

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 07:54:35 ; Search time 3376 Seconds

(without alignments)  
654.767 Million cell updates/sec

Title: US-10-032-393-36

Perfect score: 51  
Sequence: 1 tcataaaatttatttgcgt.....ttttctgataatagatca 51

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

Genbml:
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_pl:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hgt_hum:*
31: em_hgt_inv:*
32: em_hgt_other:*
33: em_hgt_mus:*
34: em_hgt_pln:*
35: em_hgt_rod:*
36: em_hgt_mam:*
37: em_hgt_vtl:*
38: em_sy:*
39: em_hgt_hum:*
40: em_hgt_mus:*
41: em_hgt_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	60	6 A15398	A15398 Synthetic P
2	51	100.0	75	6 P15PN25	M11602 Bacterioph
3	51	100.0	202	6 A10374	A10374 Nucleotide
4	51	100.0	202	6 A11246	A11246 vector pDS2
5	51	100.0	202	6 E00936	E00936 DNA sequenc
6	51	100.0	207	6 A10375	A10375 Nucleotide
7	51	100.0	247	6 A11242	A11242 vector pDS1
8	51	100.0	259	6 A10373	A10373 Nucleotide
9	51	100.0	1113	6 A02739	A02739 Artificial
10	51	100.0	1113	6 A14592	A14592 Synthetic n
11	51	100.0	1221	6 A02296	A02296 Plasmid pDS
12	51	100.0	1221	6 A14593	A14593 Synthetic n
13	51	100.0	1866	6 A02227	A02227 Plasmid pDS
14	51	100.0	1866	6 A02295	A02295 Plasmid pDS
15	51	100.0	1866	6 A02740	A02740 Artificial
16	51	100.0	1866	6 A07053	A07053 Nucleotide
17	51	100.0	1866	6 A14594	A14594 Synthetic n
18	51	100.0	3403	6 A02228	A02228 Plasmid pDS
19	51	100.0	5767	6 AX035965	AX035965 Sequence
20	51	100.0	6000	12 U66308	U66308 Expression
21	51	100.0	6447	12 AF288421	AF288421 Synthetic
22	50	98.0	76	6 A11173	A11173 A TS promot
23	46.4	91.0	524	12 SYNIFNGS	K01699 Human immu
24	46	90.2	73	6 A12013	A12013 oligonucleo
25	46	90.2	73	6 A12014	A12014 oligonucleo
26	42.8	83.9	83	6 A16630	A16630 Nucleotide
27	42.8	83.9	83	6 A16631	A16631 Nucleotide
28	39	76.5	68	6 E01976	E01976 DNA encodin
29	33.6	65.9	131	6 A50146	A50146 Sequence 3
30	33.4	65.5	125	6 A46760	A46760 Sequence 2
31	33.4	65.5	125	6 I89341	I89341 Sequence 2
32	33	64.7	116	6 A13088	A13088 Nucleotide
33	33	64.7	132	12 SYNADRE1B1	M6661 Synthetic a
34	33	64.7	171	6 AR069843	AR069843 Sequence
35	33	64.7	173	6 A01257	A01257 Nucleotide
36	33	64.7	173	6 A01258	A01258 (reverse co
37	33	64.7	173	6 A13104	A13104 Nucleotide
38	33	64.7	173	6 A13105	A13105 Nucleotide
39	33	64.7	180	6 A01247	A01247 Nucleotide
40	33	64.7	180	6 A01248	A01248 (reverse co
41	33	64.7	180	6 A13094	A13094 Nucleotide
42	33	64.7	180	6 A13095	A13095 Nucleotide
43	33	64.7	319	12 ARLACI	X06035 Synthetic 1
44	33	64.7	396	6 A11819	A11819 Artificial
45	33	64.7	396	6 A11820	A11820 Artificial

#### ALIGNMENTS

```

RESULT 1
LOCUS      A15398      60 bp      DNA
DEFINITION Synthetic Phase T5 promoter P-N 25.
ACCESSION  A15398
VERSION    A15398.1 GI:1247805
KEYWORDS
ORGANISM  synthetic construct
SOURCE    synthetic construct
          artificial sequences.
REFERENCE 1 (bases 1 to 60)
AUTHORS   Bujard,H. and Le Grice,S.
TITLE     New gram-positive expression-control sequences
          Patent EP 0207459-A2 07-JAN-1997?
JOURNAL   F. HOFFMANN-LA ROCHE AG

```

## FEATURES

Location/Qualifiers  
1..60  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"

## ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 60;  
Best Local Similarity 100.0%; Pred. No. 0.0089;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
1 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51

## RESULT 2

PTSPN25 75 bp DNA linear PHG 28-APR-1993  
DEFINITION Bacteriophage T5 promoter P-N 25.  
ACCESSION M11602  
VERSION M11602.1 GI:215985  
KEYWORDS Promoter region.  
SOURCE Bacteriophage T5  
ORGANISM Bacteriophage T5

## REFERENCE

1 (bases 1 to 75)  
Gentz, R. and Bujard, H.  
Promoters recognized by Escherichia coli RNA polymerase selected by  
function: highly efficient promoters from bacteriophage T5

## JOURNAL

J. Bacteriol. 164 (1), 70-77 (1985)

## MEDLINE

86008105

## COMMENT

Original source text: Bacteriophage T5 DNA.

## FEATURES

Location/Qualifiers  
1..75  
/organism="Bacteriophage T5"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10726"

## misc\_RNA

/note="bacteriophage T5 RNA"  
Undetermined.

## ORIGIN

Query Match 100.0%; Score 51; DB 7; Length 75;  
Best Local Similarity 100.0%; Pred. No. 0.0085;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
6 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 56

## RESULT 3

LOCUS A10374 202 bp DNA linear PAT 22-SEP-1993  
DEFINITION Nucleotide sequence 21 from patent number EP0303925.  
ACCESSION A10374  
VERSION A10374.1 GI:490696  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
1 (bases 1 to 202)  
REFERENCE Bujard, H. and Lianzer, M.  
AUTHORS High repressible sequence for control of expression  
TITLE Patent: EP 0303925-A 21 22-FEB-1989;  
JOURNAL F. HOFFMANN-LA ROCHE AG

## FEATURES

Location/Qualifiers  
1..202  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"

## ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 202;  
Best Local Similarity 100.0%; Pred. No. 0.0068;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
124 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 174

## RESULT 4

LOCUS A11246 202 bp DNA linear PAT 12-NOV-1993  
DEFINITION vector pDS2/EN25X/0.102+ XhoI/EcoRI fragment.  
ACCESSION A11246  
VERSION A11246.1 GI:491020  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.

## REFERENCE

1 (bases 1 to 202)  
Bujard, H. and Scudiero, D.  
New expression control sequence  
Patent: EP 0186069-A 10 02-JUL-1986;  
JOURNAL F. HOFFMANN-LA ROCHE AG

## FEATURES

Location/Qualifiers  
1..202  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"

## ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 202;  
Best Local Similarity 100.0%; Pred. No. 0.0068;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
124 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 174

## RESULT 5

LOCUS E00936 202 bp DNA linear PAT 29-SEP-1997  
DEFINITION DNA sequence including promoter and operator hybrid 'P(N25X/)',  
ACCESSION E00936  
VERSION E00936.1 GI:2169197  
KEYWORDS JP 1986181386-A/1.  
SOURCE Escherichia coli  
ORGANISM Escherichia coli

## REFERENCE

1 (bases 1 to 202)  
Heruman, B. and Deietoritsuhi, S.  
NOVEL DEVELOPMENT REGULATING ARRANGEMENT  
Patent: JP 1986181386-A 1 14-AUG-1986;  
JOURNAL F. HOFFMANN LA ROCHE & CO AG

## COMMENT

OS Escherichia coli  
PN JP 1986181386-A/1  
PD 14-AUG-1986  
PF 16-DEC-1985 JP 1985282699  
PR 17-DEC-1984 GB 84 8431818  
PI HERUMAN BUYARDO, DEIETORITSUHI SUTERYUBAA  
PC C12N15/00, C12N1/20, C12N9/10, C12P21/02, (C12N1/20, C12R1.19), PC  
(C12N1/20,

PC C12R1.125), (C12N9/10, C12R1.19), (C12N9/10, C12R1.125), (C12P21/02, PC  
C12R1.19),  
PC (C12P21/02, C12R1.125);  
CC strandedness: Single;  
CC topology: linear;  
CC hypothetical: No;

```
CC anti-sense: No;
CC *source: strain=M15;
FH Key Location/Qualifiers
FH
FT promoter 1..<173
FT sig_peptide /note='P25 promoter'
FT sig_peptide 174..>190
FT /note='lac operator'.

FEATURES
source
1..202 Location/Qualifiers
/mol_type='Bacterichia coli'
/mol_type='genomic DNA'
/db_xref='taxon:562'

ORIGIN
Query Match 100.0%; Score 51; DB 6; Length 202;
Best Local Similarity 100.0%; Pred. No. 0.0068;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTAATAGATTCA 51
124 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTAATAGATTCA 174

RESULT 6
LOCUS A10375 207 bp DNA linear PAT 22-SEP-1993
DEFINITION Nucleotide sequence 22 from patent number EP0303925.
ACCESSION A10375
VERSION A10375.1 GI:490697
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 207)
AUTHORS Bujard,H. and Lanzer,M.
TITLE High repressible sequence for control of expression
JOURNAL Patent: EP 0303925-A 22 22-FEB-1989;
F. HOFFMANN-LA ROCHE AG

FEATURES
source
1..207 Location/Qualifiers
/mol_type='synthetic construct'
/mol_type='unassigned DNA'
/db_xref='taxon:32630'

ORIGIN
Query Match 100.0%; Score 51; DB 6; Length 207;
Best Local Similarity 100.0%; Pred. No. 0.0068;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTAATAGATTCA 51
124 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTAATAGATTCA 174

RESULT 7
LOCUS A11242 247 bp DNA linear PAT 12-NOV-1993
DEFINITION vector pDS1/PN25 XhoI-fragment carrying promoter Pn25 is displayed.
ACCESSION A11242
VERSION A11242.1 GI:491016
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 247)
AUTHORS Bujard,H. and Stueber,D.
TITLE New expression control sequence
JOURNAL Patent: EP 0186069-A 6 02-JUL-1986;
F. HOFFMANN-LA ROCHE AG

FEATURES
source
1..247 Location/Qualifiers
/mol_type='synthetic construct'

ORIGIN
Query Match 100.0%; Score 51; DB 6; Length 247;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTAATAGATTCA 51
124 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTAATAGATTCA 174

RESULT 8
LOCUS A10373 259 bp DNA linear PAT 22-SEP-1993
DEFINITION Nucleotide sequence 20 from patent number EP0303925.
ACCESSION A10373
VERSION A10373.1 GI:490695
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 259)
AUTHORS Bujard,H. and Lanzer,M.
TITLE High repressible sequence for control of expression
JOURNAL Patent: EP 0303925-A 20 22-FEB-1989;
F. HOFFMANN-LA ROCHE AG

FEATURES
source
1..259 Location/Qualifiers
/mol_type='synthetic construct'
/mol_type='unassigned DNA'
/db_xref='taxon:32630'

ORIGIN
Query Match 100.0%; Score 51; DB 6; Length 259;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTAATAGATTCA 51
130 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTAATAGATTCA 180

RESULT 9
LOCUS A02739 1113 bp DNA linear PAT 23-MAR-1993
DEFINITION Artificial sequence of plasmid pDS5/RBSII,3A+5A (XhoI/XbaI
fragment).
ACCESSION A02739
VERSION A02739.1 GI:345272
KEYWORDS chloramphenicol acetyltransferase.
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 1113)
AUTHORS Genz,R., le Grice,S., Mous,J. and Stueber,D.
TITLE ENV/GAG polyepitides
JOURNAL Patent: EP 0270114-A 5 08-JUN-1988;
F. HOFFMANN-LA ROCHE AG

FEATURES
source
1..1113 Location/Qualifiers
/mol_type='synthetic construct'
/mol_type='unassigned DNA'
/db_xref='taxon:32630'
225..947
/codon_start=1
/transl_table=1
/product='chloramphenicol acetyltransferase'
/protein_id='CA00265.1'
/db_xref='GI:345273'
/db_xref='RENTREMBL:CA00265'
/translation='VRGSDVLPPLAFLAFSGAKAEKKEKLTGYTVDSIQWHRKEHPE'

CDS
1..247
```

AFQSVACQCTNCTGCTTGTCTTACGAGAAATTTTCTGTATATAGATTCA 51  
 LVIMDSVHPCTVTFEHEOTETSSLSMESEHDPROPLHTSYDVCYGNLAVFPGFI  
 ENMFVSNPMTVSFTSPDLNVANMNFAPAVFTMGKYYTGQDKVLMPLAIGVHNAVCD  
 GFHVGRMLNEIQCYCDEWOGGA"

## ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 1113;  
 Best Local Similarity 100.0%; Pred. No. 0.0047;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTATTTGCTTTCAGAGAAATTTTCTGTATATAGATTCA 51  
 DB 125 TCATATAAAATTTATTTGCTTTCAGAGAAATTTTCTGTATATAGATTCA 175

## RESULT 10

LOCUS A14592 1113 bp DNA circular PAT 21-FEB-1994  
 DEFINITION Synthetic nucleotide sequence encoding chloramphenicol  
 acetyltransferase (plus 21 additional N-terminus amino acids).

ACCESSION A14592  
 VERSION A14592.1 GI:491830  
 KEYWORDS synthetic construct  
 SOURCE synthetic construct  
 ORGANISM artificial sequences.

REFERENCE 1 (bases 1 to 1113)  
 Banmarch, W., Cerra, U., Mous, J. and Stueber, D.  
 Polypeptides that elicit antibodies against AIDS virus  
 Patent: EP 0219106-A 5 22-APR-1987;  
 F. HOFFMANN-LA ROCHE AG

## FEATURES

source 1..1113  
 Location/Qualifiers  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 225..947  
 /codon\_start=1  
 /transl\_table=11  
 /product="chloramphenicol acetyltransferase"  
 /protein\_id="CA01171.1"  
 /db\_xref="GI:491831"  
 /db\_xref="RENTREMBL:CA01171"  
 /translation="MRGSVDLPSTLAPSGAKEMKKITGYTTVIDISQNRKENE  
 AFQSVACQCTNCTGCTTGTCTTACGAGAAATTTTCTGTATATAGATTCA 51  
 LVIMDSVHPCTVTFEHEOTETSSLSMESEHDPROPLHTSYDVCYGNLAVFPGFI  
 ENMFVSNPMTVSFTSPDLNVANMNFAPAVFTMGKYYTGQDKVLMPLAIGVHNAVCD  
 GFHVGRMLNEIQCYCDEWOGGA"

## ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 1113;  
 Best Local Similarity 100.0%; Pred. No. 0.0047;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTATTTGCTTTCAGAGAAATTTTCTGTATATAGATTCA 51  
 DB 125 TCATATAAAATTTATTTGCTTTCAGAGAAATTTTCTGTATATAGATTCA 175

## RESULT 11

LOCUS A02296 1221 bp DNA linear PAT 18-MAY-1993  
 DEFINITION Plasmid pDS6/RBSII, 3A+5A DNA XhoI/XbaI-fragment.  
 ACCESSION A02296  
 VERSION A02296.1 GI:345259  
 KEYWORDS synthetic construct  
 SOURCE synthetic construct  
 ORGANISM artificial sequences.

REFERENCE 1 (bases 1 to 1221)  
 Cerra, U., Gentz, R. and Takacs, B. antigen peptides  
 Plasmodium falciparum mezozoite  
 Patent: EP 0283825-A 17 28-SEP-1988;  
 JOURNAL

F. HOFFMANN-LA ROCHE AG  
 Location/Qualifiers  
 1..1221  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"

## ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 1221;  
 Best Local Similarity 100.0%; Pred. No. 0.0046;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTATTTGCTTTCAGAGAAATTTTCTGTATATAGATTCA 51  
 DB 125 TCATATAAAATTTATTTGCTTTCAGAGAAATTTTCTGTATATAGATTCA 175

## RESULT 12

LOCUS A14593 1221 bp DNA circular PAT 21-FEB-1994  
 DEFINITION Synthetic nucleotide sequence of the XhoI/XbaI fragment of  
 pDS6/RBSII, 3A+5A.

ACCESSION A14593  
 VERSION A14593.1 GI:491832  
 KEYWORDS synthetic construct  
 SOURCE synthetic construct  
 ORGANISM artificial sequences.

REFERENCE 1 (bases 1 to 1221)  
 Banmarch, W., Cerra, U., Mous, J. and Stueber, D.  
 Polypeptides that elicit antibodies against AIDS virus  
 Patent: EP 0219106-A 6 22-APR-1987;  
 F. HOFFMANN-LA ROCHE AG

## FEATURES

source 1..1221  
 Location/Qualifiers  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 225..269  
 /gene="cat"  
 /gene="cat"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CA01172.1"  
 /db\_xref="GI:491833"  
 /db\_xref="RENTREMBL:CA01172"  
 /translation="MRGSVDLPSTLAPSGAKEMKKITGYTTVIDISQNRKENE  
 AFQSVACQCTNCTGCTTGTCTTACGAGAAATTTTCTGTATATAGATTCA 51  
 LVIMDSVHPCTVTFEHEOTETSSLSMESEHDPROPLHTSYDVCYGNLAVFPGFI  
 ENMFVSNPMTVSFTSPDLNVANMNFAPAVFTMGKYYTGQDKVLMPLAIGVHNAVCD  
 GFHVGRMLNEIQCYCDEWOGGA"

## ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 1221;  
 Best Local Similarity 100.0%; Pred. No. 0.0046;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTATTTGCTTTCAGAGAAATTTTCTGTATATAGATTCA 51  
 DB 125 TCATATAAAATTTATTTGCTTTCAGAGAAATTTTCTGTATATAGATTCA 175

## RESULT 13

LOCUS A02227 1866 bp DNA linear PAT 29-APR-1996  
 DEFINITION Plasmid pDS6/RBSII, Sphi DNA for XhoI/XbaI fragment.  
 ACCESSION A02227  
 VERSION A02227.1 GI:490286  
 KEYWORDS synthetic construct  
 SOURCE synthetic construct  
 ORGANISM artificial sequences.

REFERENCE 1 (bases 1 to 1866)  
 Doebeli, H., Eggmann, B., Gentz, R., Hochuli, E. and Stueber, D.  
 Fusion proteins and their purification  
 Patent: EP 0282042-A 1 14-SEP-1988;  
 JOURNAL

FEATURES F. HOFFMANN-LA ROCHE AG  
 Location/Qualifiers  
 source 1..1866  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"

Query Match 100.0%; Score 51; DB 6; Length 1866;  
 Best Local Similarity 100.0%; Pred. No. 0.0042;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAATTTATTTGCTTCAGAAATTTTCGTATATAGATTCA 51  
 |||||  
 DB 125 TCATATAAATTTATTTGCTTCAGAAATTTTCGTATATAGATTCA 175

RESULT 14  
 A02295 1866 bp DNA linear PAT 18-MAY-1993  
 LOCUS A02295  
 DEFINITION Plasmid pDS8/RBSII, SphI DNA XhoI/XbaI-fragment.  
 ACCESSION A02295  
 VERSION A02295.1 GI:345258  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 artificial sequences.

REFERENCE 1 (bases 1 to 1866)  
 AUTHORS Certz,U., Gentz,R. and Takacs,B.  
 TITLE Plasmodium falciparum merozoite antigen peptides  
 JOURNAL Patent: EP 0283829-A 16 28-SEP-1988;  
 F. HOFFMANN-LA ROCHE AG

FEATURES  
 source 1..1866  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"

## ORIGIN

Query Match 100.0%; Score 51; DB 6; Length 1866;  
 Best Local Similarity 100.0%; Pred. No. 0.0042;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAATTTATTTGCTTCAGAAATTTTCGTATATAGATTCA 51  
 |||||  
 DB 125 TCATATAAATTTATTTGCTTCAGAAATTTTCGTATATAGATTCA 175

## RESULT 15

A02740 1866 bp DNA linear PAT 23-MAR-1993  
 LOCUS A02740  
 DEFINITION Artificial sequence of pDS8/RBSII plasmid (XhoI/XbaI fragment).  
 ACCESSION A02740  
 VERSION A02740.1 GI:345274  
 KEYWORDS dihydrofolate reductase.  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 artificial sequences.

REFERENCE 1 (bases 1 to 1866)  
 AUTHORS Gentz,R., Le Grice,S., Mous,J. and Stueber,D.  
 TITLE ENY/GAG polypeptides  
 JOURNAL Patent: EP 0270114-A 6 08-JUN-1988;  
 F. HOFFMANN-LA ROCHE AG

FEATURES  
 source 1..1866  
 Location/Qualifiers

/organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 225..806  
 /codon\_start=1  
 /transl\_table=1  
 /product="dihydrofolatreductase"  
 /protein\_id="CAA00266.1"

/db\_xref="GI:345275"  
 /db\_xref="REMBL:CAA00266"  
 /translation="MRSGIMVRLPINCIVASOMGIGKNDLPPPIRNFKYFOR  
 TTSSVBGRQNLVIMGRKTPSIPKRPPLKDRINITYLSRLKXPPRGARHLASLD  
 ALRLIQPELASKVDMVIVGSSVVOEAMNQPHLRLFVTRIMQEFSSDTFPEIDL  
 GKYLIPFVPLSEVQEQEKIKKXFEYERKD"

Query Match 100.0%; Score 51; DB 6; Length 1866;  
 Best Local Similarity 100.0%; Pred. No. 0.0042;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAATTTATTTGCTTCAGAAATTTTCGTATATAGATTCA 51  
 |||||  
 DB 125 TCATATAAATTTATTTGCTTCAGAAATTTTCGTATATAGATTCA 175

Search completed: May 4, 2004, 11:34:14  
 Job time : 3379 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 07:51:37 ; Search time 389 Seconds  
(without alignments)  
556,962 Million cell updates/sec

Title: US-10-032-393-36  
Perfect score: 1 tcataaaatttatttgc.....tttctgtataaagattca 51  
Sequence: 51

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001as:\*  
5: Geneseq2001bs:\*  
6: Geneseq2002s:\*  
7: Geneseq2003as:\*  
8: Geneseq2003bs:\*  
9: Geneseq2003cs:\*  
10: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	51	6	ABK98620 Gram posi
2	51	100.0	51	8	ACD13871 Xyl-T5 pr
3	51	100.0	60	1	AAN70573 Coliphage
4	51	100.0	72	4	AAN74868 Oligonuc
5	51	100.0	80	4	AAN74869 Oligonuc
6	51	100.0	80	6	ABK98610 Gram posi
7	51	100.0	80	6	ACD13861 Xyl-T5 fu
8	51	100.0	94	6	ABK98589 Xyl-T5 f
9	51	100.0	94	8	ACD13840 Oligonuc
10	51	100.0	95	6	ABK98590 Xyl-T5 f
11	51	100.0	95	8	ACD13841 Oligonuc
12	51	100.0	200	1	AAN91066 Control s
13	51	100.0	202	1	AAN60262 Plasmid p
14	51	100.0	207	1	AAN91067 Control s
15	51	100.0	247	1	AAN60259 Coliphage
16	51	100.0	259	1	AAN91065 Control s
17	51	100.0	556	6	ABK98585 Gram posi
18	51	100.0	556	8	ACD13836 Xyl-T5 fu
19	51	100.0	1113	1	AAN80481 XhoI/XbaI
20	51	100.0	1221	1	AAN81153 XhoI/XbaI
21	51	100.0	1866	1	AAN80482 XhoI/XbaI
22	51	100.0	1866	1	AAN81356 Sequence
23	51	100.0	1866	1	AAN81152 XhoI/XbaI

24	51	100.0	1866	1	AAN80955 XhoI-XbaI
25	51	100.0	3403	1	AAN80956 Plasmid p
26	51	100.0	5302	3	AAN98012 Expressio
27	51	100.0	5767	5	AAN98014 Expressio
28	51	100.0	6852	6	ABK98600 Vector px
29	51	100.0	6852	8	ACD13851 Xyl opera
30	50	98.0	1246	4	AAN50517 Staphyloc
31	46	90.2	77	1	AAN70113 Sequence
32	43	84.3	51	6	ABK98621 Gram posi
33	43	84.3	51	8	ACD13872 Xyl-T5 mu
34	43	84.3	80	6	ABK98611 Gram posi
35	43	84.3	80	8	ACD13862 Xyl-T5 mu
36	43	84.3	94	8	ABK98606 Xyl-T5 DD
37	43	84.3	94	8	ACD13857 Xyl-T5 mu
38	43	84.3	95	6	ABK98607 Xyl-T5 DD
39	43	84.3	6852	6	ABK98637 Vector px
40	43	84.3	6852	8	ACD13888 Xyl-T5 mu
41	42.8	83.9	83	1	AAN81195 Beta-lact
42	39.4	77.3	80	1	AAN80462 Expressio
43	39	76.5	68	1	AAN81260 Hybrid pr
44	36.4	71.4	41	1	AAN90998 Fragments
45	33.4	65.5	125	2	AAT04945 Fragment

## ALIGNMENTS

RESULT 1	ABK98620	standard; DNA, 51 BP.
ID	ABK98620	
XX		
AC	ABK98620;	
XX		
DT	07-AUG-2003 (revised)	
DT	21-OCT-2002 (first entry)	
XX		
DE	Gram positive bacteria Xyl-T5 fusion promoter #3.	
XX		
KM	ds; promoter; gram positive bacteria; fusion promoter; T5; CP25; P32;	
KM	P59; P1P2; PL; XylO; tecto; trpo; malo; lambdaclO; cellular proliferation;	
KM	antibiotic.	
XX		
OS	Eubacteria.	
OS	Bacteriophage T5.	
XX		
OS	Synthetic.	
XX		
PN	WO200251982-A2	
XX		
PD	04-JUL-2002.	
XX		
PF	21-DEC-2001; 2001WO-US050250.	
XX		
XX	27-DEC-2000; 2000US-0259434P.	
PR	06-SEP-2001; 2001US-00948993.	
XX		
PA	(E)ITRA PHARM INC.	
XX		
PI	Haselbeck R, Wall D, Gross M;	
DR	WPI; 2002-575374/61.	
XX		
PT	Isolated nucleic acid comprises bacterial promoters modified to have	
PT	altered activity in at least one gram-positive organism, e.g. Bacillus	
PT	anthracis or Clostridium botulinum, useful for regulating gene expression	
PT	in bacteria.	
XX		
PS	Claim 2; Page 219; 246pp; English.	
XX		
CC	The invention relates to an isolated nucleic acid comprising a fusion	
CC	promoter comprising at least one promoter that is modified to have	
CC	altered activity in at least one gram-positive organism, or comprising	
CC	T5, CP25, P32, P59, P1P2 or PL linked to at least one operator consisting	
CC	of xylO, tecto, trpo, malo or lambdaclO, where at least one operator is	

CC positioned so binding of a repressor to an operator represses  
 CC transcription from the fusion promoter. Also included are vectors and  
 CC host cells comprising the fusion promoters, a method of identifying genes  
 CC involved in cellular proliferation or required for proliferation of a  
 CC prokaryotic cell using the vector, a method of identifying compounds that  
 CC inhibit the proliferation of a prokaryotic cell using the vector, a  
 CC method of identifying a compound that reduces the activity or level of a  
 CC gene product required for proliferation of a cell using the vector, a  
 CC compound identified by the methods, a method of inhibiting the activity  
 CC or expression of a gene in an operon required for proliferation using the  
 CC vector, manufacturing an antibiotic comprising using the vector or cell  
 CC and identifying a nucleic acid with promoter activity in *Enterococcus*  
 CC faecalis. The fusion promoters are useful for regulating nucleic acid or  
 CC polypeptide expression, particularly for regulating gene expression in  
 CC bacteria and for identifying proliferation-regulated genes or molecules  
 CC with potential antibiotic activity. The modified promoters are also  
 CC useful for replacing endogenous promoters to create cells with specific  
 CC regulatable genes. The present sequence is a fusion promoter sequence of  
 CC the invention. (Updated on 07-AUG-2003 to correct OS field.)  
 CC  
 XX

SQ Sequence 51 BP, 19 A, 5 C, 5 G, 22 T, 0 U, 0 Other;

Query Match 100.0%; Score 51; DB 6; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 2,1e-05;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCATATAAAATTATTTGCTTCAGGAAATTTTCTGTATATGATTC A 51  
 DB 1 TCATATAAAATTATTTGCTTCAGGAAATTTTCTGTATATGATTC A 51

# RESULT 2

ID ACD13871 standard; DNA; 51 BP.

AC ACD13871;

DT 15-AUG-2003 (first entry)

DE Xyl-T5 promoter sequence.

KM Promoter; ds; gram positive bacteria; *Staphylococcus aureus*;

KM *Enterococcus faecalis*; operator; xylO; tetO; trpO; malO; lambda-c10;

OS Bacteriophage T5.

XX Unidentified.

PN US2003027286-A1.

PD 06-FEB-2003.

PF 21-DEC-2001; 2001US-00032393.

PR 06-SEP-2000; 2000US-0230335P.

PA 27-DEC-2000; 2000US-0259434P.

XX (HASE/) HASELBECK R.

PA (WALL/) WALL D.

XX (GROSS/) GROSS M.

PI Haseelbeck R, Wall D, Gross M;

DR WPI; 2003-479541/45.

XX New isolated nucleic acid comprising a fusion promoter having at least

PT one promoter that is modified to have altered activity in at least one

PT gram-positive organism, useful for regulating gene expression in

PT bacteria.

PS Claim 2; Page 80; 142pp; English.

XX The invention relates to an isolated nucleic acid comprising a fusion

CC promoter having at least one promoter that is modified to have altered  
 CC activity in at least one gram-positive organism (e.g. *Staphylococcus*  
 CC *aureus* or *Enterococcus faecalis*). The promoter is linked to at least one  
 CC operator selected from xylO, tetO, trpO, malO and lambda-c10, which are  
 CC positioned such that the binding of at least one repressor to the  
 CC operator represses transcription from the fusion promoter. Also included  
 CC are vectors comprising the isolated nucleic acid, a host cell comprising  
 CC the nucleic acid. The fusion promoter is useful for identifying genes  
 CC involved in cellular proliferation, identifying a compound that reduces  
 CC the activity or level of a gene product required for proliferation of a  
 CC cell, inhibiting the activity or expression of a gene in an operon  
 CC required for proliferation, manufacturing an antibiotic, identifying a  
 CC gene that is required for proliferation of a prokaryotic cell,  
 CC identifying a compound that inhibits the proliferation of a prokaryotic  
 CC cell and regulating gene expression in bacteria. The present sequence is  
 CC a bacterial promoter suitable for inclusion in a fusion promoter of the  
 CC invention  
 CC  
 XX

SQ Sequence 51 BP, 19 A, 5 C, 5 G, 22 T, 0 U, 0 Other;

Query Match 100.0%; Score 51; DB 8; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 2,1e-05;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCATATAAAATTATTTGCTTCAGGAAATTTTCTGTATATGATTC A 51  
 DB 1 TCATATAAAATTATTTGCTTCAGGAAATTTTCTGTATATGATTC A 51

# RESULT 3

ID AAN70573 standard; DNA; 60 BP.

AC AAN70573;

DT 25-MAR-2003 (revised)

DT 11-MAR-1991 (first entry)

DE Coliphage T5 promoter of *Bacillus subtilis* expression vector.

KM Expression vector; dihydrofolate reductase, IL-2; insulin; HIV-III; tPA;

XX human renin; ds.

OS Coliphage T5.

XX Key

FT -35\_signal

FT -10\_signal

PN BP207459-A.

PD 07-JAN-1987.

PF 27-JUN-1986; 86EP-00108774.

PR 05-JUL-1985; 85GB-00017071.

PA (HOFF) HOFFMANN-LA ROCHE AG.

XX Bujard H, Legrice S;

PI WPI; 1987-001183/01.

XX New gram positive expression control DNA sequences - useful for potent

PT and versatile gene expression of prokaryotic or eukaryotic proteins in

PT *Bacillus subtilis* etc.

XX Disclosure; Table I; 53pp; English.

XX The gene fragment may be incorporated into a novel expression vector, for

CC controlling protein production in transformed Gram-positive bacteria esp.

CC Bacillus subtilis. Proteins which may be encoded include: dihydrofolate  
CC reductase; chloramphenicol acetyltransferase; malaria surface antigens;  
CC IL-2; interferons; insulin; CPA; human renin and HIV-III. (Updated on 25  
CC -MAR-2003 to correct PI field.)

XX SQ Sequence 60 BP; 23 A; 5 C; 6 G; 26 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 1; Length 60;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATTAATAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
|||||  
Db 1 TCATTAATAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51

RESULT 4  
AAH74868  
ID AAH74868 standard; DNA; 72 BP.

XX AC AAH74868;  
XX DT 29-OCT-2001 (first entry)  
XX DE

Oligonucleotide used to create T5 promoter.

XX Recombinase; genetic engineering; monogenic disorder; ADA deficiency;  
XX cystic fibrosis; familial hypercholesterolaemia; haemophilia;  
XX chronic granulomatous disease; Duchenne's muscular dystrophy;  
XX Fanconi's disease; anemia; sickle-cell anemia; Gaucher's disease;  
XX Hunter's syndrome; X-linked severe combined immunodeficiency; SCID;  
XX infectious disease; acquired disorder; tumour; cancer; T5 promoter; ss.  
XX OS Synthetic.  
XX PN WO200161049-A1.  
XX PD 23-AUG-2001.  
XX PF 16-FEB-2001; 2001WO-US005269.  
XX PR 18-FEB-2000; 2000US-0183759P.  
XX PA (STRD ) UNIV LEIAND STANFORD JUNIOR.  
XX PI Calos MP, Scilment CR;  
XX DR WPI; 2001-522610/57.  
XX DT

PT Identifying altered recombinase, involves hybridizing recombination sites  
PT in cells having specific coding sequence, transformed with altered  
PT recombinase gene, and isolating cells having product of the sequence.  
XX PS Disclosure; Page 59; 101pp; English.

XX The specification describes a method for identifying altered  
CC recombinases. The method comprises transforming cells with a first  
CC plasmid comprising two recombination sites and a coding sequence of  
CC interest, and a second plasmid encoding an altered recombinase; allowing  
CC recombination of the two recombination sites using the altered  
CC recombinase; screening and isolating transformed cells comprising the  
CC product of the sequence of interest; and identifying the altered  
CC recombinase. The altered recombinase is useful for site-specifically  
CC integrating a polynucleotide sequence of interest in a genome of a cell.  
CC It is also useful in genetic engineering of chromosomes of higher cells,  
CC and for the generation of transgenic cells, tissues, plants and animals.  
CC The altered recombinase is useful for treating monogenic disorders, e.g.  
CC ADA deficiency, cystic fibrosis, familial hypercholesterolaemia, anemia,  
CC chronic granulomatous disease, Duchenne's muscular dystrophy,  
CC haemophilia, Fanconi's disease, sickle-cell anemia, Gaucher's disease,  
CC Hunter's syndrome and X-linked severe combined immunodeficiency (SCID),  
CC infectious diseases including viral and bacterial infections, acquired  
CC disorders including solid tumours and haematopoietic cancers such as

CC leukaemias and lymphomas, and other cancers. Oligonucleotides AAH74868-69  
CC were used to create the T5 promoter, which was used to construct a  
CC plasmid for use in the course of the invention

XX SQ Sequence 72 BP; 27 A; 6 C; 11 G; 28 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 4; Length 72;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATTAATAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
|||||  
Db 2 TCATTAATAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 52

RESULT 5  
AAH74869/C  
ID AAH74869 standard; DNA; 80 BP.

XX AC AAH74869;  
XX DT 29-OCT-2001 (first entry)  
XX DE

Oligonucleotide used to create T5 promoter.

XX Recombinase; genetic engineering; monogenic disorder; ADA deficiency;  
XX cystic fibrosis; familial hypercholesterolaemia; haemophilia;  
XX chronic granulomatous disease; Duchenne's muscular dystrophy;  
XX Fanconi's disease; anemia; sickle-cell anemia; Gaucher's disease;  
XX Hunter's syndrome; X-linked severe combined immunodeficiency; SCID;  
XX infectious disease; acquired disorder; tumour; cancer; T5 promoter; ss.  
XX OS Synthetic.  
XX PN WO200161049-A1.  
XX PD 23-AUG-2001.  
XX PF 16-FEB-2001; 2001WO-US005269.  
XX PR 18-FEB-2000; 2000US-0183759P.  
XX PA (STRD ) UNIV LEIAND STANFORD JUNIOR.  
XX PI Calos MP, Scilment CR;  
XX DR WPI; 2001-522610/57.  
XX DT

PT Identifying altered recombinase, involves hybridizing recombination sites  
PT in cells having specific coding sequence, transformed with altered  
PT recombinase gene, and isolating cells having product of the sequence.  
XX PS Disclosure; Page 59; 101pp; English.

XX The specification describes a method for identifying altered  
CC recombinases. The method comprises transforming cells with a first  
CC plasmid comprising two recombination sites and a coding sequence of  
CC interest, and a second plasmid encoding an altered recombinase; allowing  
CC recombination of the two recombination sites using the altered  
CC recombinase; screening and isolating transformed cells comprising the  
CC product of the sequence of interest; and identifying the altered  
CC recombinase. The altered recombinase is useful for site-specifically  
CC integrating a polynucleotide sequence of interest in a genome of a cell.  
CC It is also useful in genetic engineering of chromosomes of higher cells,  
CC and for the generation of transgenic cells, tissues, plants and animals.  
CC The altered recombinase is useful for treating monogenic disorders, e.g.  
CC ADA deficiency, cystic fibrosis, familial hypercholesterolaemia, anemia,  
CC chronic granulomatous disease, Duchenne's muscular dystrophy,  
CC haemophilia, Fanconi's disease, sickle-cell anemia, Gaucher's disease,  
CC Hunter's syndrome and X-linked severe combined immunodeficiency (SCID),  
CC infectious diseases including viral and bacterial infections, acquired  
CC disorders including solid tumours and haematopoietic cancers such as  
CC leukaemias and lymphomas, and other cancers. Oligonucleotides AAH74868-69



CC were used to create the T5 promoter, which was used to construct a  
CC plasmid for use in the course of the invention

XX Sequence 80 BP; 29 A; 14 C; 9 G; 28 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 4; Length 80;

Best Local Similarity 100.0%; Pred. No. 2.1e-05;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTATTGCTTCAGGAAATTTTCGTATATAGATTCA 51

Db 75 TCATATAAAATTATTGCTTCAGGAAATTTTCGTATATAGATTCA 25

#### RESULT 6

ABK98610  
ID ABK98610 standard; DNA; 80 BP.

XX ABK98610;

AC 07-AUG-2003 (revised)

DT 21-OCT-2002 (first entry)

XX Gram positive bacteria Xyl-T5 fusion promoter #2.

DE ds; promoter; gram positive bacteria; fusion promoter; T5; CP25; P32;

KM P59; P1P2; PL; xylO; tetO; trpO; malO; lambdaclO; cellular proliferation;

KW antibiotic.

XX Eubacteria.

OS Bacteriophage T5.

OS Synthetic.

PN WO200251982-A2.

XX 04-JUL-2002.

PD 21-DEC-2001; 2001WO-US050250.

PF 27-DEC-2000; 2000US-0259434P.

XX 06-SEP-2001; 2001US-00948993.

PR (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Wall D, Gross M;

PI WPI; 2002-575374/61.

XX Isolated nucleic acid comprises bacterial promoters modified to have

PT altered activity in at least one gram-positive organism, e.g. *Bacillus*

PT *anthracis* or *Clostridium botulinum*, useful for regulating gene expression

PT in bacteria.

XX Claim 24; Page 217; 246pp; English.

XX The invention relates to an isolated nucleic acid comprising a fusion

CC promoter comprising at least one promoter that is modified to have

CC altered activity in at least one gram-positive organism, or comprising

CC T5, CP25, P32, P59, P1P2 or PL linked to at least one operator consisting

CC of xylO, tetO, trpO, malO or lambdaclO, where at least one operator is

CC positioned so binding of a repressor to an operator represses

CC transcription from the fusion promoter. Also included are vectors and

CC host cells comprising the fusion promoters. A method of identifying genes

CC involved in cellular proliferation or required for proliferation of a

CC prokaryotic cell using the vector, a method of identifying compounds that

CC inhibit the proliferation of a prokaryotic cell using the vector, a

CC method of identifying a compound that reduces the activity or level of a

CC gene product required for proliferation of a cell using the vector, a

CC compound identified by the methods, a method of inhibiting the activity

CC or expression of a gene in an operon required for proliferation using the

CC vector, manufacturing an antibiotic comprising using the vector or cell

CC and identifying a nucleic acid with promoter activity in *Enterococcus*

CC *faecalis*. The fusion promoters are useful for regulating nucleic acid or

CC polypeptide expression, particularly for regulating gene expression in

CC bacteria and for identifying proliferation-regulated genes or molecules

CC with potential antibiotic activity. The modified promoters are also

CC useful for replacing endogenous promoters to create cells with specific

CC regulatable genes. The present sequence is a fusion promoter sequence of

CC the invention. (Updated on 07-AUG-2003 to correct OS field.)

XX Sequence 80 BP; 29 A; 8 C; 8 G; 35 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 6; Length 80;

Best Local Similarity 100.0%; Pred. No. 2.1e-05;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTATTGCTTCAGGAAATTTTCGTATATAGATTCA 51

Db 1 TCATATAAAATTATTGCTTCAGGAAATTTTCGTATATAGATTCA 51

#### RESULT 7

ACD13861  
ID ACD13861 standard; DNA; 80 BP.

XX ACD13861;

AC 15-AUG-2003 (first entry)

DT Xyl-T5 fusion promoter #1.

XX Promoter; ds; gram positive bacteria; *Staphylococcus aureus*;

XX *Enterococcus faecalis*; operator; xylO; tetO; trpO; malO; lambda-clO;

XX cellular proliferation.

XX Bacteriophage T5.

OS Unidentified.

PN US2003027286-A1.

XX 06-FEB-2003.

PD 21-DEC-2001; 2001US-00032393.

PF 06-SEP-2000; 2000US-0230335P.

XX 27-DEC-2000; 2000US-0259434P.

PR (HASE/) HASELBECK R.

XX (WALL/) WALL D.

PA (GROSS/) GROSS M.

PI Haselbeck R, Wall D, Gross M;

XX WPI; 2003-479541/45.

XX New isolated nucleic acid comprising a fusion promoter having at least

PT one promoter that is modified to have altered activity in at least one

PT gram-positive organism, useful for regulating gene expression in

PT bacteria.

XX Claim 24; Page 78; 142pp; English.

XX The invention relates to an isolated nucleic acid comprising a fusion

CC promoter having at least one promoter that is modified to have altered

CC activity in at least one gram-positive organism (e.g. *Staphylococcus*

CC *aureus* or *Enterococcus faecalis*). The promoter is linked to at least one

CC operator selected from xylO, tetO, trpO, malO and lambda-clO, which are

CC positioned such that the binding of at least one repressor to the

CC operator represses transcription from the fusion promoter. Also included

CC are a vector comprising the isolated nucleic acid, a host cell comprising

CC the nucleic acid. The fusion promoter is useful for identifying genes

CC involved in cellular proliferation, identifying a compound that reduces

CC the activity or level of a gene product required for proliferation of a

CC cell, inhibiting the activity or expression of a gene in an operon

CC required for proliferation, manufacturing an antibiotic, identifying a

CC gene that is required for proliferation of a prokaryotic cell.

CC identifying a compound that inhibits the proliferation of a prokaryotic  
CC cell and regulating gene expression in bacteria. The present sequence is  
CC a fusion promoter of the invention

XX Sequence 80 BP, 29 A, 8 C, 8 G, 35 T, 0 U, 0 Other;

Query Match 100.0%; Score 51; DB 8; Length 80;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAATAAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 1 TCAATAAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51

RESULT 8  
ABK98589  
ID ABK98589 standard; DNA; 94 BP.

XX ABK98589;  
AC 07-AUG-2003 (revised)  
DT 21-OCT-2002 (first entry)

DE Xy10-T5 fusion promoter oligonucleotide #1.

XX ss; promoter; gram positive bacteria; fusion promoter; T5; CP25; P32;  
KM P59; P1P2; PL; xyl0; tetO; trpO; mal0; lambdacl0; cellular proliferation;  
KM antibiotic.

XX Eubacteria.  
OS Bacteriophage T5.  
OS Synthetic.

PN WO200251982-A2.

PD 04-JUL-2002.

PF 21-DEC-2001; 2001WO-US050250.

PR 27-DEC-2000; 2000US-0259434P.

PR 06-SEP-2001; 2001US-00948993.

XX (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Wall D, Gross M;

DR WPI; 2002-575374/61.

XX Isolated nucleic acid comprises bacterial promoters modified to have  
PT altered activity in at least one gram-positive organism, e.g. *Bacillus*  
PT *anthracis* or *Clostridium botulinum*, useful for regulating gene expression  
PT in bacteria.

PS Example 1; Page 81; 246pp; English.

XX The invention relates to an isolated nucleic acid comprising a fusion  
CC promoter comprising at least one promoter that is modified to have  
CC altered activity in at least one gram-positive organism, or comprising  
CC T5, CP25, P32, P59, P1P2 or PL linked to at least one operator consisting  
CC of xyl0, tetO, trpO, mal0 or lambdacl0, where at least one operator is  
CC positioned so binding of a repressor to an operator represses  
CC transcription from the fusion promoter. Also included are vectors and  
CC host cells comprising the fusion promoters, a method of identifying genes  
CC involved in cellular proliferation or required for proliferation of a  
CC prokaryotic cell using the vector, a method of identifying compounds that  
CC inhibit the proliferation of a prokaryotic cell using the vector, a  
CC method of identifying a compound that reduces the activity or level of a  
CC gene product required for proliferation of a cell using the vector, a  
CC compound identified by the methods, a method of inhibiting the activity  
CC or expression of a gene in an operon required for proliferation using the  
CC vector, manufacturing an antibiotic comprising using the vector or cell  
CC and identifying a nucleic acid with promoter activity in *Enterococcus*

CC faecalis. The fusion promoters are useful for regulating nucleic acid or  
CC polypeptide expression, particularly for regulating gene expression in  
CC bacteria and for identifying proliferation-regulated genes or molecules  
CC with potential antibiotic activity. The modified promoters are also  
CC useful for replacing endogenous promoters to create cells with specific  
CC regulatable genes. The present sequence is an oligonucleotide used to  
CC construct a fusion promoter sequence of the invention. (Updated on 07-AUG  
CC -2003 to correct OS field.)  
XX

QY Sequence 94 BP, 35 A, 10 C, 12 G, 37 T, 0 U, 0 Other;  
Query Match 100.0%; Score 51; DB 6; Length 94;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAATAAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 6 TCAATAAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 56

RESULT 9  
ACD13840  
ID ACD13840 standard; DNA; 94 BP.

XX ACD13840;

DT 15-AUG-2003 (first entry)

DE Oligonucleotide Xyl-T5.

XX Promoter; ss; gram positive bacteria; *Staphylococcus aureus*;  
KM *Enterococcus faecalis*; operator; xyl0; tetO; trpO; mal0; lambda-cl0;  
KM cellular proliferation.

XX *Escherichia coli*.  
OS Synthetic.  
OS Unidentified.

PN US2003027286-A1.

PD 06-FEB-2003.

PF 21-DEC-2001; 2001US-00032393.

PR 06-SEP-2000; 2000US-0230335P.

PR 27-DEC-2000; 2000US-0259434P.

XX (HASE/) HASELBECK R.

PA (WALL/) WALL D.

PI (GROSS/) GROSS M.

DR Haselbeck R, Wall D, Gross M;

DR WPI; 2003-479541/45.

XX New isolated nucleic acid comprising a fusion promoter having at least  
PT one promoter that is modified to have altered activity in at least one  
PT gram-positive organism, useful for regulating gene expression in  
PT bacteria.

PS Example 1; Page 28; 142pp; English.

XX The invention relates to an isolated nucleic acid comprising a fusion  
CC promoter having at least one promoter that is modified to have altered  
CC activity in at least one gram-positive organism (e.g. *Staphylococcus*  
CC *aureus* or *Enterococcus faecalis*). The promoter is linked to at least one  
CC operator or selected from xyl0, tetO, trpO, mal0 and lambda-cl0, which are  
CC positioned such that the binding of at least one repressor to the  
CC operator represses transcription from the fusion promoter. Also included  
CC are a vector comprising the isolated nucleic acid, a host cell comprising  
CC the nucleic acid. The fusion promoter is useful for identifying genes  
CC involved in cellular proliferation, identifying a compound that reduces  
CC the activity or level of a gene product required for proliferation of a

CC cell, inhibiting the activity or expression of a gene in an operon  
CC required for proliferation, manufacturing an antibiotic, identifying a  
CC gene that is required for proliferation of a prokaryotic cell,  
CC identifying a compound that inhibits the proliferation of a prokaryotic  
CC cell and regulating gene expression in bacteria. The present sequence is  
CC an oligonucleotide used in the construction of a fusion promoter of the  
CC invention.

XX Sequence 94 BP; 35 A; 10 C; 12 G; 37 T; 0 U; 0 Other;

Query Match	100.0%	Score 51	DB 8	Length 94
Best Local Similarity	100.0%	Pred. No. 2	1e-05	
Matches 51	Conservative 0	Mismatches 0	Indels 0	Gaps 0

```

Qy      1 TCATATAAAATTTATTTGCTTTCAGGAAATTTTCGTATAATGATTTCA 51
        |||||
Db      6 TCATATAAAATTTATTTGCTTTCAGGAAATTTTCGTATAATGATTTCA 56

```

RESULT 10  
ABK98590/c  
ID ABK98590 standard; DNA; 95 BP.

DT	07-AUG-2003	(revised)
DT	21-OCT-2002	(first entry)

CC are a vector comprising the isolated nucleic acid, a host cell comprising  
CC the nucleic acid. The fusion promoter is useful for identifying genes  
CC involved in cellular proliferation, identifying a compound that reduces  
CC the activity or level of a gene product required for proliferation of a  
CC cell, inhibiting the activity or expression of a gene in an operon  
CC required for proliferation, manufacturing an antibiotic, identifying a  
CC gene that is required for proliferation of a prokaryotic cell,  
CC identifying a compound that inhibits the proliferation of a prokaryotic  
CC cell and regulating gene expression in bacteria. The present sequence is  
CC an oligonucleotide used in the construction of a fusion promoter of the  
XX invention

XX SQ Sequence 95 BP; 39 A; 11 C; 9 G; 36 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 8; Length 95;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATTAATAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 93 TCATTAATAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 43

## RESULT 12

AAN91066  
ID AAN91066 standard; DNA; 200 BP.

XX AAN91066;

XX AC 03-OCT-2002 (revised)  
XX DT 09-JUL-1989 (first entry)

XX DE Control sequence N25\*/O.

XX KM DNA expression control; ss.

XX OS Synthetic.

XX EN EP303925-A.

XX PD 22-FEB-1989.

XX PF 08-AUG-1988; 88EP-00112864.

XX FR 17-AUG-1987; 87CH-00003152.

XX PA (HOFF) HOFFMANN-LA ROCHE AG.

XX PI Bujard H, Lanzer M;

XX DR WPI; 1989-055375/08.

XX PT Deoxyribonucleic acid data expression control sequences - comprising  
XX promoter and operator-repressor-sequences.

XX PS Disclosure; Fig 7; 65pp; German.

XX CC This sequence is the operator N25\*/O and is inserted into plasmid pDS3  
XX for expression of a variety of proteins from pro- and eukaryotic sources.  
XX It is prepared as a DNA XhoI/EcoRI frag. This operator/repressor sequence  
XX has a high complexing rate and gives good repressibility. See also  
XX CC N910601-3,5,7,8 and AAN91070. (Updated on 03-OCT-2002 to add missing OS  
XX field.)  
XX SQ Sequence 200 BP; 64 A; 33 C; 37 G; 66 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 1; Length 200;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATTAATAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 124 TCATTAATAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 174

## RESULT 13

AAN60262  
ID AAN60262 standard; DNA; 202 BP.

XX AAN60262;

XX AC 29-MAY-1991 (first entry)

XX DT Plasmid pDS2/P promoter/operator fusion (PN25x/O).

XX DE DHFR; chloramphenicol acetyl-transferase; malaria SA; IL-2; IFN; insulin;  
XX KM tPA; renin; ds.

XX OS Synthetic.

XX PN EP16069-A.

XX PD 02-JUL-1986.

XX PF 13-DEC-1985; 88EP-00115921.

XX PR 17-DEC-1984; 84GB-00031818.

XX PA (HOFF) HOFFMANN-LA ROCHE AG.

XX PI Bujard H, Stuber D;

XX DR WPI; 1986-170629/27.

XX PT Expression control DNA sequence - comprising 15 promoter combined with  
XX DNA sequence which permits control of promoter activity.

XX PS Disclosure; Fig 7; 26pp; English.

XX CC Plasmid vectors of the pDS1 family may be used to express a sequence  
XX under the control of the coliphage T5 promoter, and one or more  
XX CC sequences which allow control of the promoter. Sequences expressed  
XX include products such as dihydrofolate reductase; chloramphenicol acetyl-  
XX transferase; malaria surface antigen; IL-2; IFN-alpha, -beta and -gamma;  
XX CC insulin; growth hormones; tPA; human renin etc

XX SQ Sequence 202 BP; 64 A; 34 C; 37 G; 67 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 1; Length 202;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATTAATAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 124 TCATTAATAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 174

## RESULT 14

AAN91067  
ID AAN91067 standard; DNA; 207 BP.

XX AAN91067;

XX DT 03-OCT-2002 (revised)

XX DT 09-JUL-1989 (first entry)

XX DE Control sequence N25OP29.

XX KM DNA expression control; ss.

XX OS Synthetic.

XX PN EP303925-A.

XX PD 22-FEB-1989.

PF 08-AUG-1988; 88EP-00112864.  
 XX  
 PR 17-AUG-1987; 87CH-00003152.  
 XX  
 PA (HOFF) HOFFMANN-LA ROCHE AG.  
 XX  
 PI Bujard H, Lanzer M;  
 XX  
 DR WPI; 1989-055375/08.  
 XX  
 PT Deoxyribonucleic acid data expression control sequences - comprising  
 PS Promoter and operator-repressor sequences.  
 XX  
 PS Disclosure; Fig 7; 65pp; German.  
 XX  
 CC Sequence is the operator/promoter N250p29 and is inserted into plasmid  
 CC pDS3 for expression of a variety of proteins from pro- and eukaryotic  
 CC sources. It is prepared as a DNA XhoI/EcoRI frag. It comprises a low  
 CC signal strength/high promoter strength promoter and high complexing rate  
 CC operator/repressor. It gives high transcription and translation efficien-  
 CC ies and good repressability. See also AAN91060, 1,2,3,5,6,8 and AAN91070.  
 CC (Updated on 03-OCT-2002 to add missing OS field.)  
 CC  
 SQ Sequence 207 BP; 66 A; 37 C; 37 G; 67 T; 0 U; 0 Other;  
 XX  
 Query Match 100.0%; Score 51; DB 1; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 2,1e-05;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 TCATATAAAATTATTTCCTTCAGGAAATTTTCGTATATAGATTCA 51  
 DB 124 TCATATAAAATTATTTCCTTCAGGAAATTTTCGTATATAGATTCA 174

RESULT 15  
 AAN60259  
 ID AAN60259 standard; DNA; 247 BP.  
 XX  
 AC AAN60259;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 29-MAY-1991 (first entry)  
 XX  
 DE Coliphage PN25 Pre-early promoter of plasmid pDSL, PN25, col+.  
 XX  
 KM DHFR; chloramphenicol acetyl-transferase; malaria SA; IL-2; IFN; insulin;  
 KM tPA; renin.  
 XX  
 OS Unidentified.  
 OS  
 XX  
 PN EPI86069-A.  
 XX  
 PD 02-JUL-1986.  
 XX  
 PF 13-DEC-1985; 85EP-00115921.  
 XX  
 PR 17-DEC-1984; 84GB-00031818.  
 XX  
 PA (HOFF) HOFFMANN-LA ROCHE AG.  
 XX  
 PI Bujard H, Studer D;  
 XX  
 DR WPI; 1986-170629/27.  
 XX  
 PT Expression control DNA sequence - comprising T5 promoter combined with  
 PT DNA sequence which permits control of promoter activity.  
 XX  
 PS Disclosure; Fig 2; 26pp; English.  
 XX  
 CC Plasmid vectors of the pSL family may be used to express a sequence  
 CC under the control of a coliphage T5 promoter eg. the coliphage PN25 pre-  
 CC early promoter, and one or more sequences which allow control of the  
 CC promoter. Sequences expressed include products such as dihydrofolate

CC reductase; chloramphenicol acetyl-transferase; malaria surface antigen;  
 CC IL-2; IFN-alpha, -beta and -gamma; insulin; growth hormones; tPA; human  
 CC renin etc. (Updated on 27-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 247 BP; 79 A; 44 C; 48 G; 76 T; 0 U; 0 Other;  
 XX  
 Query Match 100.0%; Score 51; DB 1; Length 247;  
 Best Local Similarity 100.0%; Pred. No. 2,1e-05;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 TCATATAAAATTATTTCCTTCAGGAAATTTTCGTATATAGATTCA 51  
 DB 124 TCATATAAAATTATTTCCTTCAGGAAATTTTCGTATATAGATTCA 174

Search completed: May 4, 2004, 10:37:46  
 Job time : 391 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 09:55:29 ; Search time 76 Seconds  
(without alignments)  
372.401 Million cell updates/sec

Title: US-10-032-393-36  
Perfect score: 51  
Sequence: 1 tcatataaaattcttgcct.....ttctctgataatagattca 51

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33.4	65.5	125	1	US-08-400-864-2 Sequence 2, Appl
2	33	64.7	171	2	US-08-892-272-3 Sequence 3, Appl
3	33	64.7	757	2	US-08-892-272-1 Sequence 1, Appl
4	31.4	61.6	49	1	US-08-400-864-1 Sequence 1, Appl
5	28.6	56.1	247	3	US-09-344-888A-24 Sequence 2, Appl
6	28.6	56.1	250	3	US-08-840-466A-25 Sequence 2, Appl
7	28.6	56.1	250	4	US-09-696-188B-25 Sequence 25, Appl
8	28.6	56.1	520	1	US-08-268-348A-7 Sequence 7, Appl
9	28.6	56.1	520	1	US-08-268-348A-9 Sequence 9, Appl
10	28.6	56.1	1106	1	US-08-041-648-4 Sequence 4, Appl
11	28.6	56.1	1106	1	US-08-041-648-6 Sequence 6, Appl
12	28.6	56.1	1106	1	US-08-041-648-8 Sequence 8, Appl
13	28.6	56.1	3462	4	US-07-794-400-2 Sequence 10, Appl
14	28.6	56.1	3977	1	US-07-794-400-2 Sequence 2, Appl
15	28.6	56.1	3977	1	US-07-794-400-13 Sequence 1, Appl
16	28.6	56.1	3977	1	US-08-041-648-1 Sequence 1, Appl
17	28.6	56.1	3977	1	US-08-217-529-1 Sequence 1, Appl
18	28.6	56.1	3977	1	US-08-397-470-2 Sequence 2, Appl
19	28.6	56.1	3977	1	US-08-397-470-13 Sequence 13, Appl
20	28.6	56.1	4202	4	US-09-483-419-2 Sequence 2, Appl
21	28.6	56.1	4491	4	US-09-837-863-23 Sequence 23, Appl
22	28.6	56.1	4755	4	US-09-837-863-24 Sequence 24, Appl
23	28.6	56.1	6501	4	US-09-767-515-1 Sequence 1, Appl
24	28.6	56.1	6501	4	US-09-767-515-2 Sequence 2, Appl
25	27.2	53.3	1830121	4	US-09-557-884-1 Sequence 1, Appl
26	27.2	53.3	1830121	4	US-09-643-990A-1 Sequence 1, Appl
27	27	52.9	4519	4	US-08-976-259-6 Sequence 6, Appl

C	28	25.4	49.8	843	4	US-09-134-001C-2669	Sequence 2669, Ap
	29	25	49.0	92407	4	US-09-596-002-36	Sequence 36, Appl
	30	25	49.0	640681	4	US-09-790-968-1	Sequence 1, Appl
C	31	24.8	48.6	6152	3	US-08-973-462-1	Sequence 1, Appl
C	32	24.4	47.8	2728	4	US-09-620-312D-572	Sequence 572, App
C	33	24.4	47.8	15788	4	US-09-920-759-13	Sequence 13, Appl
	34	24.4	47.8	193303	4	US-09-497-855A-37	Sequence 37, Appl
	35	24.4	47.8	193303	4	US-09-497-855A-44	Sequence 44, Appl
	36	24.2	47.5	1650	4	US-09-907-794A-254	Sequence 254, App
	37	24.2	47.5	1650	4	US-09-905-125A-254	Sequence 254, App
C	38	24.2	47.5	1650	4	US-09-902-775A-254	Sequence 254, App
	39	24	47.1	640681	4	US-09-790-968-1	Sequence 1, Appl
	40	23.8	46.7	786431	4	US-09-751-389-3	Sequence 1, Appl
	41	23.6	46.3	2243	1	US-07-995-657-1	Sequence 1, Appl
	42	23.6	46.3	2243	1	US-08-474-587-1	Sequence 1, Appl
	43	23.6	46.3	6113	2	US-08-974-546-2	Sequence 2, Appl
C	44	23.6	46.3	6113	4	US-10-204-708-14	Sequence 14, Appl
	45	23.6	46.3	112132	4	US-09-741-150-3	Sequence 3, Appl

## ALIGNMENTS

RESULT 1  
US-08-400-864-2  
; Sequence 2, Application US/08400864  
; Patent No. 5721137  
; GENERAL INFORMATION:  
; APPLICANT: PRASCOITI, GIANNI  
; APPLICANT: GRANDI, GUIDO  
; TITLE OF INVENTION: PLASMID VECTOR AND ITS USE FOR THE  
; TITLE OF INVENTION: PRODUCTION OF HETEROLOGOUS PROTEINS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSSTADT  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/400,864  
; FILING DATE: 08-MAR-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: IT 000727 MI94A  
; FILING DATE: 15-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 2264-083-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 125 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-400-864-2  
Query Match 65.5%; Score 33.4; DB 1; Length 125;  
Best Local Similarity 95.7%; Pred. No. 0.16;  
Matches 45; Conservative 0; Mismatches 1; Indels 1;

QY 6 AAAATTTATTTGCTTCAGAGAAA-TTTTCTGTATATATAGATCA 51  
Db 5 AAAATTTATTTGCTTCAGAGAAA-TTTTATGTATATATAGATCA 51

## RESULT 2

US-08-892-272-3  
Sequence 3, Application US/08892272  
Patent No. 5891687  
GENERAL INFORMATION:  
APPLICANT: SCHLIEPER, Daniel  
APPLICANT: SOBER, Harald  
APPLICANT: SCHMIDT, Manfred  
APPLICANT: VON WILKEN-BERGEMANN, Brigitte  
APPLICANT: MULLER-HILL, Benno  
TITLE OF INVENTION: POSITIVE SELECTION VECTOR BASED ON THE  
TITLE OF INVENTION: CAPS GENE, PCAPS VECTOR AND ITS USE  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nikolaide, Marmelstein, Murray & Oram  
STREET: 655 Fifteenth Street N.W. Suite 330  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/892,272  
FILING DATE: 14-JUL-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 196 30 617.5  
FILING DATE: 29-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Murray, Robert B.  
REGISTRATION NUMBER: 22,980  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)638-5000  
TELEFAX: (202)638-4810  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 171 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-892-272-3

Query Match 64.7%; Score 33; DB 2; Length 171;  
Best Local Similarity 79.6%; Pred. No. 0.21;  
Matches 39; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 ATAAAAATTTATTTGCTTCAGAGAAA-TTTTCTGTATATATAGATCA 51  
Db 71 ATAAAAATTTATTTGCTTCAGAGATCTTGATATATATATATATCA 119

## RESULT 3

US-08-892-272-1  
Sequence 1, Application US/08892272  
Patent No. 5891687  
GENERAL INFORMATION:  
APPLICANT: SCHLIEPER, Daniel  
APPLICANT: SOBER, Harald  
APPLICANT: SCHMIDT, Manfred  
APPLICANT: VON WILKEN-BERGEMANN, Brigitte  
APPLICANT: MULLER-HILL, Benno  
TITLE OF INVENTION: POSITIVE SELECTION VECTOR BASED ON THE  
TITLE OF INVENTION: CAPS GENE, PCAPS VECTOR AND ITS USE

NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nikolaide, Marmelstein, Murray & Oram  
STREET: 655 Fifteenth Street N.W. Suite 330  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/892,272  
FILING DATE: 14-JUL-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 196 30 617.5  
FILING DATE: 29-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Murray, Robert B.  
REGISTRATION NUMBER: 22,980  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)638-5000  
TELEFAX: (202)638-4810  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 757 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 90..722  
US-08-892-272-1

Query Match 64.7%; Score 33; DB 2; Length 757;  
Best Local Similarity 79.6%; Pred. No. 0.22;  
Matches 39; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 ATAAAAATTTATTTGCTTCAGAGAAA-TTTTCTGTATATATAGATCA 51  
Db 13 ATAAAAATTTATTTGCTTCAGAGATCTTGATATATATATATATCA 61

RESULT 4  
US-08-400-864-1  
Sequence 1, Application US/08400864  
Patent No. 5721137  
GENERAL INFORMATION:  
APPLICANT: FRASCOTTI, GIANNI  
APPLICANT: GRANDI, GUIDO  
TITLE OF INVENTION: PLASMID VECTOR AND ITS USE FOR THE  
TITLE OF INVENTION: PRODUCTION OF HETEROLOGOUS PROTEINS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/400,864  
FILING DATE: 08-MAR-1995  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IT 000727 M194A  
FILING DATE: 15-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 2264-083-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-400-864-1

Query Match 61.6%; Score 31.4; DB 1; Length 49;  
Best Local Similarity 95.6%; Pred. No. 0.57;  
Matches 43; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Cy 6 AAAAATTATTTGCTTCAGAGAAA-TTTTCTGTATATAGATT 49  
Db 5 AAAAATTATTTGCTTCAGAGAAA-TTTTATGTATATAGATT 49

RESULT 5  
US-09-344-888A-24  
Sequence 24, Application US/09344888A  
Patent No. 6291245  
GENERAL INFORMATION:  
APPLICANT: Kopeckzi, Erhard  
APPLICANT: Schantz, Christian  
TITLE OF INVENTION: New Host-Vector System  
FILE REFERENCE: CD20315  
CURRENT APPLICATION NUMBER: US/09/344, 888A  
CURRENT FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: EP98113156.8  
PRIOR FILING DATE: 1998-07-15  
PRIOR APPLICATION NUMBER: EP98119078.8  
PRIOR FILING DATE: 1998-10-09  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 24  
LENGTH: 247  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: expression cassette  
US-09-344-888A-24

Query Match 56.1%; Score 28.6; DB 3; Length 247;  
Best Local Similarity 72.5%; Pred. No. 3.8;  
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Cy 1 TCATATAAAATTTATTTGCTTCAGAGAAA-TTTTCTGTATATAGATTCA 51  
Db 10 TCATATAAAATTTATTTGCTTCAGAGGCGATACACATTTATATATAGATTCA 60

RESULT 6  
US-08-840-466A-25  
Sequence 25, Application US/08840466A  
Patent No. 6261561  
GENERAL INFORMATION:  
APPLICANT: Stewart, C. Neal  
McKee, Marian L.  
O'Brien, Alison D.  
Wachtel, Marian R.  
TITLE OF INVENTION: Method Of Stimulating An Immune Response

By Administration Of Host Organisms That Express Intimin  
Alone Or As A Fusion Protein With One Or More Other  
Antigens.  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
Dunner, L.L.P.  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/840,466A  
FILING DATE: 18-Apr-1997  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Boone, Laura L.  
REGISTRATION NUMBER: 43,505  
REFERENCE/DOCKET NUMBER: 04995-0029-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 408-4000  
TELEFAX: (202) 408-4400  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 250 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
US-08-840-466A-25

Query Match 56.1%; Score 28.6; DB 3; Length 250;  
Best Local Similarity 72.5%; Pred. No. 3.8;  
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Cy 1 TCATATAAAATTTATTTGCTTCAGAGAAA-TTTTCTGTATATAGATTCA 51  
Db 10 TCATATAAAATTTATTTGCTTCAGAGGCGATACACATTTATATATAGATTCA 60

RESULT 7  
US-09-696-188B-25  
Sequence 25, Application US/09696188B  
Patent No. 6406885  
GENERAL INFORMATION:  
APPLICANT: Stewart, C. Neal  
McKee, Marian L.  
O'Brien, Alison D.  
Wachtel, Marian R.  
TITLE OF INVENTION: Method Of Stimulating An Immune Response  
By Administration Of Host Organisms That Express Intimin  
Alone Or As A Fusion Protein With One Or More Other  
Antigens.  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
Dunner, L.L.P.  
STREET: 1300 I Street, N.W., Suite 700  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS



SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/696,188B  
FILING DATE: 26-Oct-2000  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/840,466  
FILING DATE: 1997-04-18  
ATTORNEY/AGENT INFORMATION:  
NAME: Boone, Laurel S.  
REGISTRATION NUMBER: 43,505  
REFERENCE/DOCKET NUMBER: 04995,0029-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 408-4000  
TELEFAX: (202) 408-4400  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 250 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
US-09-696-188B-25

Query Match 56.1%; Score 28.6; DB 4; Length 250;  
Best Local Similarity 72.5%; Pred. No. 3.8;  
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

CY 1 TCATATAAATTATTGCTTCAGGAAATTTTCGTATATAGATTCA 51  
|||||  
Db 10 TCATATAAATTATTGCTTCGTAGCGGATACCAATTATATATAGATTCA 60  
|||||

## RESULT 8

US-08-268-348A-7

Sequence 7, Application US/08268348A  
Patent No. 5750374  
GENERAL INFORMATION:  
APPLICANT: Dobell, Heinz  
APPLICANT: Draeger, Nicholas  
APPLICANT: Trotteman, Gerda H  
APPLICANT: Jakob, Peter  
TITLE OF INVENTION: Process for Producing Hydrophobic  
TITLE OF INVENTION: Polypeptides and Proteins, and Fusion Proteins for Use in  
TITLE OF INVENTION: Producing Same  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hoffmann-La Roche Inc.  
STREET: 340 Kingsland Street  
CITY: Nutley  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/268,348A  
FILING DATE: 29-JUN-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 93110755.1  
FILING DATE: 06-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Parise, John P.  
REGISTRATION NUMBER: 34,403  
REFERENCE/DOCKET NUMBER: 4105/157  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201) 235-6326

TELEFAX: (201) 235-3500  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 520 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 115..516  
OTHER INFORMATION: /product= "Amyloid Protein AA"  
US-08-268-348A-7

Query Match 56.1%; Score 28.6; DB 1; Length 520;  
Best Local Similarity 72.5%; Pred. No. 4;  
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

CY 1 TCATATAAATTATTGCTTCAGGAAATTTTCGTATATAGATTCA 51  
|||||  
Db 10 TCATATAAATTATTGCTTCGTAGCGGATACCAATTATATATAGATTCA 60  
|||||

## RESULT 9

US-08-268-348A-9

Sequence 9, Application US/08268348A  
Patent No. 5750374  
GENERAL INFORMATION:  
APPLICANT: Dobell, Heinz  
APPLICANT: Draeger, Nicholas  
APPLICANT: Trotteman, Gerda H  
APPLICANT: Jakob, Peter  
TITLE OF INVENTION: Process for Producing Hydrophobic  
TITLE OF INVENTION: Polypeptides and Proteins, and Fusion Proteins for Use in  
TITLE OF INVENTION: Producing Same  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hoffmann-La Roche Inc.  
STREET: 340 Kingsland Street  
CITY: Nutley  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/268,348A  
FILING DATE: 29-JUN-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 93110755.1  
FILING DATE: 06-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Parise, John P.  
REGISTRATION NUMBER: 34,403  
REFERENCE/DOCKET NUMBER: 4105/157  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201) 235-6326  
TELEFAX: (201) 235-3500  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 520 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 115..516

NAME/KEY: CDS  
LOCATION: 115..516

OTHER INFORMATION: /product= "Amyloid Protein AA"  
US-08-268-348A-9

Query Match 56.1%; Score 28.6; DB 1; Length 520;  
Best Local Similarity 72.5%; Pred. No. 4;  
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTTTCCTTCAGGAAATTTTCGTATATAGATTCA 51  
DB 10 TCATATAAAATTTTTCCTTCAGGCGGATACAAATTATATAGATTCA 60

## RESULT 10

US-08-041-648-4  
Sequence 4, Application US/08041648  
Patent No. 5486463  
GENERAL INFORMATION:  
APPLICANT: Lesslauer, Werner  
APPLICANT: L. tscher, Hansruedi  
APPLICANT: St ber, Dietrich  
TITLE OF INVENTION: TNF-MUTAINS  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.  
STREET: 340 Kingsland Street  
CITY: Nutley  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07110-1199  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/041,648  
FILING DATE: 1-APR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 92810249.0  
FILING DATE: 2-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Roseman, Catherine R.  
REGISTRATION NUMBER: 34240  
REFERENCE/DOCKET NUMBER: RAN 4105/147  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201) 235-6208  
TELEFAX: (201) 235-3500  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1106 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Plasmid pDS56/RBSII, SphI-TNFAIpha (Tnp32)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 994..1104  
US-08-041-648-4

Query Match 56.1%; Score 28.6; DB 1; Length 1106;  
Best Local Similarity 72.5%; Pred. No. 4.1; Indels 0; Gaps 0;  
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
QY 1 TCATATAAAATTTTTCCTTCAGGAAATTTTCGTATATAGATTCA 51  
DB 886 TCATATAAAATTTTTCCTTCAGGCGGATACAAATTATATAGATTCA 936

## RESULT 11

US-08-041-648-6  
Sequence 6, Application US/08041648  
Patent No. 5486463  
GENERAL INFORMATION:  
APPLICANT: Lesslauer, Werner  
APPLICANT: L. tscher, Hansruedi  
APPLICANT: St ber, Dietrich  
TITLE OF INVENTION: TNF-MUTAINS  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.  
STREET: 340 Kingsland Street  
CITY: Nutley  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07110-1199  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/041,648  
FILING DATE: 1-APR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 92810249.0  
FILING DATE: 2-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Roseman, Catherine R.  
REGISTRATION NUMBER: 34240  
REFERENCE/DOCKET NUMBER: RAN 4105/147  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201) 235-6208  
TELEFAX: (201) 235-3500  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1106 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Plasmid pDS56/RBSII, SphI-TNFAIpha (Ser29)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 994..1104  
US-08-041-648-6

Query Match 56.1%; Score 28.6; DB 1; Length 1106;  
Best Local Similarity 72.5%; Pred. No. 4.1; Indels 0; Gaps 0;  
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTTTCCTTCAGGAAATTTTCGTATATAGATTCA 51  
DB 886 TCATATAAAATTTTTCCTTCAGGCGGATACAAATTATATAGATTCA 936

## RESULT 12

US-08-041-648-8  
Sequence 8, Application US/08041648  
Patent No. 5486463  
GENERAL INFORMATION:  
APPLICANT: Lesslauer, Werner  
APPLICANT: L. tscher, Hansruedi  
APPLICANT: St ber, Dietrich  
TITLE OF INVENTION: TNF-MUTAINS  
NUMBER OF SEQUENCES: 17

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: George M. Gould, Esq., Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07110-1199

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/041.648
FILING DATE: 1-APR-1993
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: EP 92810249.0
FILING DATE: 2-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Roseman, Catherine R.
REGISTRATION NUMBER: 34240
REFERENCE/DOCKET NUMBER: RAN 4105/147
TELEPHONE: (201) 235-6208
TELEFAX: (201) 235-3500
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1106 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Plasmid pDS56/RBSII, SphI-TNFalpha(Sez29Tfp32)
FEATURE:
NAME/KEY: CDS
LOCATION: 994..1104
US-08-041-648-8

Query Match
Best Local Similarity 56.1%; Score 28.6; DB 1; Length 1106;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
DB 886 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 936

RESULT 13
US-09-742-373-10
Sequence 10, Application US/09742373
Patent No. 6562946
GENERAL INFORMATION:
APPLICANT: Althaus, Harald
APPLICANT: Hauser, Hans-Peter
TITLE OF INVENTION: Human procalcitonin and the Preparation and Use Thereof
FILE REFERENCE: 05552.1445-00
CURRENT APPLICATION NUMBER: US/09/742,373
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 19962434.8
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: 10016278.9
PRIOR FILING DATE: 2000-04-03
PRIOR APPLICATION NUMBER: 10027954.6
PRIOR FILING DATE: 2000-06-08
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 10
LENGTH: 3462
```

```

TYPE: DNA
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Vectorsequence,
US-09-742-373-10

Query Match
Best Local Similarity 56.1%; Score 28.6; DB 4; Length 3462;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
DB 10 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 60

RESULT 14
US-07-794-400-2
Sequence 2, Application US/07794400
Patent No. 5422104
GENERAL INFORMATION:
APPLICANT: Fliers, W.
APPLICANT: Tavernier, J.
APPLICANT: Van Oostede, X.
TITLE OF INVENTION: TNF-Mutlains
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: USA
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/794,400
FILING DATE: 19911120
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: EP 90810901.0
FILING DATE: 21-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Kirovatin, William
REGISTRATION NUMBER: 33256
REFERENCE/DOCKET NUMBER: 4105/136-00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-4387
TELEFAX: (201) 235-3500
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3977 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (recombinant plasmid)
IMMEDIATE SOURCE:
CLONE: pDS56/RBSII, SphI-TNF-alpha
FEATURE:
NAME/KEY: CDS
LOCATION: 115..591
US-07-794-400-2

Query Match
Best Local Similarity 56.1%; Score 28.6; DB 1; Length 3977;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
DB 10 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 60
```

```
RESULT 15
US-07-794-400-13
; Sequence 13, Application US/07794400
; Patent No. 5422104
; GENERAL INFORMATION:
; APPLICANT: Piers, W.
; APPLICANT: Tavernier, J.
; APPLICANT: Van Ostad, X.
; TITLE OF INVENTION: TNF-Mutlens
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/794,400
; FILING DATE: 1991.120
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 90810901.0
; FILING DATE: 21-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Krowatin, William
; REGISTRATION NUMBER: 33256
; REFERENCE/DOCKET NUMBER: 4105/136-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-4387
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3977 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (recombinant plasmid)
; IMMEDIATE SOURCE:
; CLONE: pDS56/RBS11, Sph1-TNF-alpha
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 115..591
; FEATURE:
; NAME/KEY: Modified site
; LOCATION: 202-204, 208-210 and 211-213
; OTHER INFORMATION: /note= "N = A, G, C or T"
US-07-794-400-13

Query Match 56.1%; Score 28.6; DB 1; Length 3977;
Best Local Similarity 72.5%; Pred. No. 4.3;
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

CY 1 TCATATAAATTATTGCTTCACGAAATTTTCTGTATATAGATTCA 51
    |||||
DB 10 TCATATAAATTATTATTGCTTCGTGACGCGATACCAATTATATAGATTCA 60
    |||||
```

Search completed: May 4, 2004, 12:20:51  
Job time : 85 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	51	100.0	51	15	US-10-032-393-36	Sequence 36, App
2	51	100.0	72	9	US-09-788-497-6	Sequence 6, App1
c	51	100.0	80	9	US-09-788-497-7	Sequence 7, App1
3	51	100.0	80	15	US-10-032-393-26	Sequence 26, App1
4	51	100.0	94	15	US-10-032-393-5	Sequence 5, App1
5	51	100.0	95	15	US-10-032-393-6	Sequence 6, App1
c	51	100.0	95	15	US-10-032-393-1	Sequence 1, App1
7	51	100.0	556	15	US-10-032-393-1	Sequence 1, App1
8	51	100.0	5302	16	US-10-985-415-2	Sequence 2, App1
9	51	100.0	5767	16	US-10-985-415-4	Sequence 4, App1
10	51	100.0	6852	15	US-10-032-393-16	Sequence 16, App
11	50	98.0	1246	9	US-09-815-242-3094	Sequence 3094, App
12	43	88.3	51	15	US-10-032-393-37	Sequence 37, App1
13	43	88.3	80	15	US-10-032-393-27	Sequence 27, App1
14	43	88.3	94	15	US-10-032-393-22	Sequence 22, App1

C	15	43	84.3	95	15	US-10-032-993-23	Sequence 23, Appl
	16	43	84.3	16852	15	US-10-032-993-33	Sequence 53, Appl
	17	29.8	58.4	113515	15	US-10-311-455-2148	Sequence 2148, Appl
	18	28.6	56.1	82	15	US-10-288-858-13	Sequence 13, Appl
	19	28.6	56.1	94	15	US-10-032-993-12	Sequence 12, Appl
	20	28.6	56.1	95	15	US-10-032-993-13	Sequence 13, Appl
C	21	28.6	56.1	134	15	US-10-284-083-3	Sequence 3, Appl
	22	28.6	56.1	138	15	US-10-284-083-4	Sequence 4, Appl
	23	28.6	56.1	138	15	US-10-284-083-5	Sequence 5, Appl
	24	28.6	56.1	162	9	US-09-828-5232A-97	Sequence 97, Appl
	25	28.6	56.1	162	10	US-09-966-621-89	Sequence 89, Appl
	26	28.6	56.1	162	15	US-10-429-094-89	Sequence 89, Appl
	27	28.6	56.1	170	15	US-10-284-083-2	Sequence 2, Appl
	28	28.6	56.1	170	15	US-10-284-083-6	Sequence 6, Appl
	29	28.6	56.1	250	8	US-08-837-459-25	Sequence 25, Appl
	30	28.6	56.1	250	15	US-10-150-058-25	Sequence 25, Appl
	31	28.6	56.1	2220	15	US-10-288-858-5	Sequence 5, Appl
	32	28.6	56.1	3420	9	US-10-385-415-1	Sequence 1, Appl
	33	28.6	56.1	3462	9	US-09-742-373-10	Sequence 10, Appl
	34	28.6	56.1	3462	15	US-10-334-058-10	Sequence 10, Appl
	35	28.6	56.1	3876	16	US-10-385-415-27	Sequence 27, Appl
	36	28.6	56.1	3876	16	US-10-385-415-28	Sequence 28, Appl
	37	28.6	56.1	3879	16	US-10-385-415-5	Sequence 5, Appl
	38	28.6	56.1	3879	16	US-10-385-415-6	Sequence 6, Appl
	39	28.6	56.1	3879	16	US-10-385-415-7	Sequence 7, Appl
	40	28.6	56.1	3879	16	US-10-385-415-21	Sequence 21, Appl
	41	28.6	56.1	3879	16	US-10-385-415-25	Sequence 25, Appl
	42	28.6	56.1	3879	16	US-10-385-415-26	Sequence 26, Appl
	43	28.6	56.1	3885	16	US-10-385-415-3	Sequence 3, Appl
	44	28.6	56.1	3887	16	US-10-385-415-20	Sequence 20, Appl
	45	28.6	56.1	3900	16	US-10-385-415-9	Sequence 9, Appl

## ALIGNMENTS

## RESULT 1

```

? Sequence 36, Application US/10032393
? Publication No. US20030027286A1
? GENERAL INFORMATION:
? APPLICANT: Haselbeck, Robert
? APPLICANT: Wall, Daniel
? APPLICANT: Gross, Molly
? TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
? FILE REFERENCE: ELITRA.010A
? CURRENT APPLICATION NUMBER: US/10/032,393
? CURRENT FILING DATE: 2001-12-21
? PRIOR APPLICATION NUMBER: 60/259,434
? PRIOR FILING DATE: 2000-12-27
? PRIOR APPLICATION NUMBER: 09/948,993
? PRIOR FILING DATE: 2001-09-06
? PRIOR APPLICATION NUMBER: 60/230,335
? PRIOR FILING DATE: 2000-09-06
? NUMBER OF SEQ ID NOS: 68
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 36
? LENGTH: 51
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Xy1-T5 promoter sequence
US-10-032-393-36

```

Query Match	100.0%;	Score 51;	DB 15;	Length 51;
Best Local Similarity	100.0%;	Pred. No. 1.5e-05;		
Matches 51; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

```
Oy      1 TCATAAAAAATTATTTGCTTCAGAAAATTTTTCTGTATAATAGATTCA   51
        |||||
Db      1 TCATAAAAAATTATTTGCCTTACGAAATTTTCTGTATAATAGATTCA   51
```

## RESULT 2

US-09-788-297-6  
Sequence 6, Application US/09788297  
Patent No. US20020094516A1  
GENERAL INFORMATION:  
APPLICANT: Calos, Michele P.  
APPLICANT: Scilmenti, Christopher R.  
TITLE OF INVENTION: ALTERED RECOMBINASES FOR GENOME MODIFICATION  
FILE REFERENCE: 8400-0011  
CURRENT APPLICATION NUMBER: US/09/788,297  
CURRENT FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 72  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: T5, oligo 1  
US-09-788-297-6

Query Match 100.0%; Score 51; DB 9; Length 72;  
Best Local Similarity 100.0%; Pred. No. 1,7e-05;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 2 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 52

## RESULT 3

US-09-788-297-7/C  
Sequence 7, Application US/09788297  
Patent No. US20020094516A1  
GENERAL INFORMATION:  
APPLICANT: Calos, Michele P.  
APPLICANT: Scilmenti, Christopher R.  
TITLE OF INVENTION: ALTERED RECOMBINASES FOR GENOME MODIFICATION  
FILE REFERENCE: 8400-0011  
CURRENT APPLICATION NUMBER: US/09/788,297  
CURRENT FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 7  
LENGTH: 80  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: T5, oligo 2  
US-09-788-297-7

Query Match 100.0%; Score 51; DB 9; Length 80;  
Best Local Similarity 100.0%; Pred. No. 1,7e-05;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 75 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 25

## RESULT 4

US-10-032-393-26  
Sequence 26, Application US/10032393  
Publication No. US20030027286A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Wall, Daniel  
APPLICANT: Gross, Molly  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010A  
CURRENT APPLICATION NUMBER: US/10/032,393  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/259,434

## PRIORITY FILING DATE: 2000-12-27

PRIOR APPLICATION NUMBER: 09/948,993  
PRIOR FILING DATE: 2001-09-06  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 26  
LENGTH: 80  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Xyl-T5 fusion sequence  
US-10-032-393-26

Query Match 100.0%; Score 51; DB 15; Length 80;  
Best Local Similarity 100.0%; Pred. No. 1,7e-05;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51

## RESULT 5

US-10-032-393-5  
Sequence 5, Application US/10032393  
Publication No. US20030027286A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Wall, Daniel  
APPLICANT: Gross, Molly  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010A  
CURRENT APPLICATION NUMBER: US/10/032,393  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/259,434  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: 09/948,993  
PRIOR FILING DATE: 2001-09-06  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 94  
TYPE: DNA  
ORGANISM: Escherichia coli  
FEATURE:  
OTHER INFORMATION: Oligonucleotide Xyl-T5  
US-10-032-393-5

Query Match 100.0%; Score 51; DB 15; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1,7e-05;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 6 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 56

## RESULT 6

US-10-032-393-6/C  
Sequence 6, Application US/10032393  
Publication No. US20030027286A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Wall, Daniel  
APPLICANT: Gross, Molly  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010A  
CURRENT APPLICATION NUMBER: US/10/032,393  
CURRENT FILING DATE: 2001-12-21

```

; PRIOR APPLICATION NUMBER: 60/259,434
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 95
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide Xyl-T5 complement
US-10-032-393-6

```

```

Query Match      100.0%; Score 51; DB 15; Length 95;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
          |||
DB      93 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 43

```

```

RESULT 7
US-10-032-393-1

```

```

; Sequence 1, Application US/10032393
; Publication No. US20030027286A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Wall, Daniel
; APPLICANT: Gross, Molly
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
; FILE REFERENCE: ELITRA.010A
; CURRENT APPLICATION NUMBER: US/10/032,393
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/259,434
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence map showing the Xyl-T5 fusion promoter
; OTHER INFORMATION: and associated DNA sequences
US-10-032-393-1

```

```

Query Match      100.0%; Score 51; DB 15; Length 556;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
          |||
DB      6 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 56

```

```

RESULT 8
US-10-385-415-2

```

```

; Sequence 2, Application US/10385415
; Publication No. US20040014158A1
; GENERAL INFORMATION:
; APPLICANT: Bacher, Adelbert
; APPLICANT: Fischer, Marius
; TITLE OF INVENTION: PROTEIN CONJUGATES, METHOD, VECTORS, PROTEINS AND DNA FOR
; TITLE OF INVENTION: PRODUCING THEM, THEIR USE AND MEDICAMENTS AND VACCINES CONTAININ
; TITLE OF INVENTION: A CERTAIN QUANTITY OF SAID PROTEIN CONJUGATES

```

```

; FILE REFERENCE: 9286.6CT
; CURRENT APPLICATION NUMBER: US/10/385,415
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: US 09/936,028
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/EP00/01899
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: DE 19910102.7
; PRIOR FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 5302
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p6021-CAT Expression vector
US-10-385-415-2

```

```

Query Match      100.0%; Score 51; DB 16; Length 5302;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
          |||
DB      134 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 184

```

```

RESULT 9
US-10-385-415-4

```

```

; Sequence 4, Application US/10385415
; Publication No. US20040014158A1
; GENERAL INFORMATION:
; APPLICANT: Bacher, Adelbert
; APPLICANT: Fischer, Marius
; TITLE OF INVENTION: PROTEIN CONJUGATES, METHOD, VECTORS, PROTEINS AND DNA FOR
; TITLE OF INVENTION: PRODUCING THEM, THEIR USE AND MEDICAMENTS AND VACCINES CONTAINI
; TITLE OF INVENTION: A CERTAIN QUANTITY OF SAID PROTEIN CONJUGATES
; FILE REFERENCE: 9286.6CT
; CURRENT APPLICATION NUMBER: US/10/385,415
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: US 09/936,028
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/EP00/01899
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: DE 19910102.7
; PRIOR FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 5767
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: p602-B5-LusY Expression plasmid
US-10-385-415-4

```

```

Query Match      100.0%; Score 51; DB 16; Length 5767;
Best Local Similarity 100.0%; Pred. No. 4.3e-05;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
          |||
DB      134 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 184

```

```

RESULT 10

```

```

US-10-032-393-16
; Sequence 16, Application US/10032393
; Publication No. US20030027286A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Wall, Daniel

```

APPLICANT: Gross, Molly  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010A  
CURRENT APPLICATION NUMBER: US/10/032,393  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/259,434  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: 09/948,993  
PRIOR FILING DATE: 2001-09-06  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 16  
LENGTH: 6852  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Vector pXyl-T5 p15a  
US-10-032-393-16

Query Match 100.0%; Score 51; DB 15; Length 6852;  
Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 TCATATAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 1367 TCATATAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 1417

RESULT 11  
US-09-815-242-3094  
Sequence 3094, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3094  
LENGTH: 1246  
TYPE: DNA  
ORGANISM: Staphylococcus aureus  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(1246)  
OTHER INFORMATION: n = A,T,C or G

Query Match 98.0%; Score 50; DB 9; Length 1246;  
Best Local Similarity 98.0%; Pred. No. 6e-05;  
Matches 50; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CY 1 TCATATAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 13 TCATATAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 63

RESULT 12  
US-10-032-393-37  
Sequence 37, Application US/10032393  
Publication No. US20030027286A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Wall, Daniel  
APPLICANT: Gross, Molly  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010A  
CURRENT APPLICATION NUMBER: US/10/032,393  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/259,434  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: 09/948,993  
PRIOR FILING DATE: 2001-09-06  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 37  
LENGTH: 51  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Xyl-T5-DD promoter sequence  
US-10-032-393-37

Query Match 84.3%; Score 43; DB 15; Length 51;  
Best Local Similarity 90.2%; Pred. No. 0.003;  
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

CY 1 TCATATAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 1 TCATATAAATTATTTTACATCAGGAAATTTTCTGTATATAGATTCA 51

RESULT 13  
US-10-032-393-27  
Sequence 27, Application US/10032393  
Publication No. US20030027286A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Wall, Daniel  
APPLICANT: Gross, Molly  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010A  
CURRENT APPLICATION NUMBER: US/10/032,393  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/259,434  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: 09/948,993  
PRIOR FILING DATE: 2001-09-06  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 27  
LENGTH: 80  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Xyl-T5-DD fusion sequence



US-10-032-393-27

Query Match 84.3%; Score 43; DB 15; Length 80;  
Best Local Similarity 90.2%; Pred. No. 0.0034;  
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCATAAAATTTATTGCTTCAGGAAATTTTCGTATATAGATTCA 51  
DB 1 TCATAAAATTTATTTCATCATCAGGAAATTTTCGTATATAGATTCA 51

RESULT 14

US-10-032-393-22  
Sequence 22, Application US/10032393  
Publication No. US20030027286A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Mail, Daniel  
APPLICANT: Gross, Molly  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010A  
CURRENT APPLICATION NUMBER: US/10/032,393  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/259,434  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: 09/948,993  
PRIOR FILING DATE: 2001-09-06  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 22  
LENGTH: 94  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Oligonucleotide Xyl-T5-DD  
US-10-032-393-22

Query Match 84.3%; Score 43; DB 15; Length 94;  
Best Local Similarity 90.2%; Pred. No. 0.0035;  
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCATAAAATTTATTGCTTCAGGAAATTTTCGTATATAGATTCA 51  
DB 6 TCATAAAATTTATTTCATCATCAGGAAATTTTCGTATATAGATTCA 56

RESULT 15

US-10-032-393-23/c  
Sequence 23, Application US/10032393  
Publication No. US20030027286A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Mail, Daniel  
APPLICANT: Gross, Molly  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010A  
CURRENT APPLICATION NUMBER: US/10/032,393  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/259,434  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: 09/948,993  
PRIOR FILING DATE: 2001-09-06  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 23  
LENGTH: 95  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:

OTHER INFORMATION: Oligonucleotide Xyl-T5-DD complement  
US-10-032-393-23

Query Match 84.3%; Score 43; DB 15; Length 95;  
Best Local Similarity 90.2%; Pred. No. 0.0035;  
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCATAAAATTTATTGCTTCAGGAAATTTTCGTATATAGATTCA 51  
DB 93 TCATAAAATTTATTTCATCATCAGGAAATTTTCGTATATAGATTCA 43

Search completed: May 4, 2004, 13:24:16  
Job time : 405 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 10:31:40 ; Search time 3247 Seconds  
(without alignments)  
552.662 Million cell updates/sec

Title: US-10-032-393-36

Perfect score: 51  
Sequence: 1 tcaataaaattatttctgct.....ttctctataatagattca 51

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 37577330 segs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents NA Main:\*

- 1: /cgn2\_6/ptodata/2/pna/PCTUS\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/pna/US08\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq:\*
- 8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pna/US092\_COMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pna/US092A\_COMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pna/US093A\_COMB.seq:\*
- 19: /cgn2\_6/ptodata/2/pna/US093B\_COMB.seq:\*
- 20: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq:\*
- 21: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq:\*
- 22: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq:\*
- 23: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq:\*
- 24: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq:\*
- 25: /cgn2\_6/ptodata/2/pna/US096A\_COMB.seq:\*
- 26: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq:\*
- 27: /cgn2\_6/ptodata/2/pna/US096C\_COMB.seq:\*
- 28: /cgn2\_6/ptodata/2/pna/US096D\_COMB.seq:\*
- 29: /cgn2\_6/ptodata/2/pna/US096E\_COMB.seq:\*
- 30: /cgn2\_6/ptodata/2/pna/US097A\_COMB.seq:\*
- 31: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq:\*
- 32: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq:\*
- 33: /cgn2\_6/ptodata/2/pna/US098A\_COMB.seq:\*
- 34: /cgn2\_6/ptodata/2/pna/US098B\_COMB.seq:\*
- 35: /cgn2\_6/ptodata/2/pna/US098C\_COMB.seq:\*
- 36: /cgn2\_6/ptodata/2/pna/US098D\_COMB.seq:\*
- 37: /cgn2\_6/ptodata/2/pna/US099A\_COMB.seq:\*
- 38: /cgn2\_6/ptodata/2/pna/US099B\_COMB.seq:\*
- 39: /cgn2\_6/ptodata/2/pna/US099C\_COMB.seq:\*
- 40: /cgn2\_6/ptodata/2/pna/US099D\_COMB.seq:\*
- 41: /cgn2\_6/ptodata/2/pna/US099E\_COMB.seq:\*
- 42: /cgn2\_6/ptodata/2/pna/US099F\_COMB.seq:\*
- 43: /cgn2\_6/ptodata/2/pna/US100A\_COMB.seq:\*

- 44: /cgn2\_6/ptodata/2/pna/US100B\_COMB.seq:\*
- 45: /cgn2\_6/ptodata/2/pna/US101A\_COMB.seq:\*
- 46: /cgn2\_6/ptodata/2/pna/US101B\_COMB.seq:\*
- 47: /cgn2\_6/ptodata/2/pna/US102A\_COMB.seq:\*
- 48: /cgn2\_6/ptodata/2/pna/US102B\_COMB.seq:\*
- 49: /cgn2\_6/ptodata/2/pna/US103A\_COMB.seq:\*
- 50: /cgn2\_6/ptodata/2/pna/US103B\_COMB.seq:\*
- 51: /cgn2\_6/ptodata/2/pna/US104A\_COMB.seq:\*
- 52: /cgn2\_6/ptodata/2/pna/US104B\_COMB.seq:\*
- 53: /cgn2\_6/ptodata/2/pna/US105A\_COMB.seq:\*
- 54: /cgn2\_6/ptodata/2/pna/US107A\_COMB.seq:\*
- 55: /cgn2\_6/ptodata/2/pna/US107B\_COMB.seq:\*
- 56: /cgn2\_6/ptodata/2/pna/US6000\_COMB.seq:\*
- 57: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq:\*
- 58: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq:\*
- 59: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq:\*
- 60: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq:\*
- 61: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq:\*
- 62: /cgn2\_6/ptodata/2/pna/US6006\_COMB.seq:\*
- 63: /cgn2\_6/ptodata/2/pna/US6007\_COMB.seq:\*
- 64: /cgn2\_6/ptodata/2/pna/US6008\_COMB.seq:\*
- 65: /cgn2\_6/ptodata/2/pna/US6009\_COMB.seq:\*
- 66: /cgn2\_6/ptodata/2/pna/US6010\_COMB.seq:\*
- 67: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq:\*
- 68: /cgn2\_6/ptodata/2/pna/US6012\_COMB.seq:\*
- 69: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq:\*
- 70: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq:\*
- 71: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq:\*
- 72: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq:\*
- 73: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq:\*
- 74: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq:\*
- 75: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq:\*
- 76: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq:\*
- 77: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq:\*
- 78: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq:\*
- 79: /cgn2\_6/ptodata/2/pna/US6023\_COMB.seq:\*
- 80: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq:\*
- 81: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq:\*
- 82: /cgn2\_6/ptodata/2/pna/US6026\_COMB.seq:\*
- 83: /cgn2\_6/ptodata/2/pna/US6027\_COMB.seq:\*
- 84: /cgn2\_6/ptodata/2/pna/US6028\_COMB.seq:\*
- 85: /cgn2\_6/ptodata/2/pna/US6029\_COMB.seq:\*
- 86: /cgn2\_6/ptodata/2/pna/US6030\_COMB.seq:\*
- 87: /cgn2\_6/ptodata/2/pna/US6031\_COMB.seq:\*
- 88: /cgn2\_6/ptodata/2/pna/US6032\_COMB.seq:\*
- 89: /cgn2\_6/ptodata/2/pna/US6033\_COMB.seq:\*
- 90: /cgn2\_6/ptodata/2/pna/US6034\_COMB.seq:\*
- 91: /cgn2\_6/ptodata/2/pna/US6035\_COMB.seq:\*
- 92: /cgn2\_6/ptodata/2/pna/US6036\_COMB.seq:\*
- 93: /cgn2\_6/ptodata/2/pna/US6037\_COMB.seq:\*
- 94: /cgn2\_6/ptodata/2/pna/US6038\_COMB.seq:\*
- 95: /cgn2\_6/ptodata/2/pna/US6039\_COMB.seq:\*
- 96: /cgn2\_6/ptodata/2/pna/US6040\_COMB.seq:\*
- 97: /cgn2\_6/ptodata/2/pna/US6041\_COMB.seq:\*
- 98: /cgn2\_6/ptodata/2/pna/US6042\_COMB.seq:\*
- 99: /cgn2\_6/ptodata/2/pna/US6043\_COMB.seq:\*
- 100: /cgn2\_6/ptodata/2/pna/US6044\_COMB.seq:\*
- 101: /cgn2\_6/ptodata/2/pna/US6045\_COMB.seq:\*
- 102: /cgn2\_6/ptodata/2/pna/US6046\_COMB.seq:\*
- 103: /cgn2\_6/ptodata/2/pna/US6047\_COMB.seq:\*
- 104: /cgn2\_6/ptodata/2/pna/US6048\_COMB.seq:\*
- 105: /cgn2\_6/ptodata/2/pna/US6049\_COMB.seq:\*
- 106: /cgn2\_6/ptodata/2/pna/US6050\_COMB.seq:\*
- 107: /cgn2\_6/ptodata/2/pna/US6051\_COMB.seq:\*
- 108: /cgn2\_6/ptodata/2/pna/US6052\_COMB.seq:\*
- 109: /cgn2\_6/ptodata/2/pna/US6053\_COMB.seq:\*
- 110: /cgn2\_6/ptodata/2/pna/US6054\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score being printed, and is derived by analysts of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	51	1	PCT-US01-50250-36
2	51	100.0	51	1	PCT-US01-50250A-36
3	51	100.0	51	43	US-10-032-393-36
4	51	100.0	72	1	PCT-US01-05269-6
5	51	100.0	72	32	US-09-788-297-6
6	51	100.0	80	1	PCT-US01-05269-7
7	51	100.0	80	1	PCT-US01-50250-26
8	51	100.0	80	1	PCT-US01-50250A-26
9	51	100.0	80	32	US-09-788-297-7
10	51	100.0	80	43	US-10-032-393-26
11	51	100.0	94	1	PCT-US01-50250-5
12	51	100.0	94	1	PCT-US01-50250A-5
13	51	100.0	94	43	US-10-032-393-5
14	51	100.0	94	82	US-60-259-434-3
15	51	100.0	95	1	PCT-US01-50250-6
16	51	100.0	95	1	PCT-US01-50250A-6
17	51	100.0	95	43	US-10-032-393-6
18	51	100.0	95	82	US-60-259-434-4
19	51	100.0	265	76	US-60-206-848-99
20	51	100.0	297	76	US-60-206-848-22
21	51	100.0	343	82	US-60-253-625-1609
22	51	100.0	384	76	US-60-206-848-232
23	51	100.0	556	1	PCT-US01-50250-1
24	51	100.0	556	1	PCT-US01-50250A-1
25	51	100.0	556	43	US-10-032-393-1
26	51	100.0	556	82	US-60-259-434-10
27	51	100.0	5302	38	US-09-936-028-2
28	51	100.0	5302	38	US-09-936-028A-2
29	51	100.0	5302	50	US-10-385-415-2
30	51	100.0	5767	38	US-09-936-028-4
31	51	100.0	5767	50	US-09-936-028A-4
32	51	100.0	5767	50	US-10-385-415-4
33	51	100.0	6852	1	PCT-US01-50250-16
34	51	100.0	6852	1	PCT-US01-50250A-16
35	51	100.0	6852	43	US-10-032-393-16
36	51	100.0	6858	82	US-60-259-434-8
37	50	98.0	1246	1	PCT-US02-03987-3094
38	50	98.0	1246	33	US-09-815-242-3094
39	50	98.0	1246	44	US-10-072-851-3094
40	44	86.3	44	1	PCT-US03-12044-15
41	44	86.3	44	1	PCT-US03-12044A-15
42	43	84.3	51	1	PCT-US01-50250-37
43	43	84.3	51	1	PCT-US01-50250A-37
44	43	84.3	51	43	US-10-032-393-37
45	43	84.3	80	1	PCT-US01-50250-27

## ALIGNMENTS

RESULT 1  
 PCT-US01-50250-36  
 ; Sequence 36, Application PC/TUS0150250  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ELITRA PHARMACEUTICALS, INC.  
 ; APPLICANT: HASELBECK, Robert  
 ; APPLICANT: WALL, Daniel  
 ; APPLICANT: GROSS, Molly  
 ; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
 ; FILE REFERENCE: ELITRA.010VPC  
 ; CURRENT APPLICATION NUMBER: PCT/US01/50250  
 ; CURRENT FILING DATE: 2001-12-21  
 ; PRIOR APPLICATION NUMBER: US 60/259,434  
 ; PRIOR FILING DATE: 2000-12-27  
 ; PRIOR APPLICATION NUMBER: US 09/948,993  
 ; PRIOR FILING DATE: 2001-09-06  
 ; NUMBER OF SEQ ID NOS: 68  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 36

```

; LENGTH: 51
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Xyl-T5 promoter sequence
PCT-US01-50250-36

Query Match      100.0%; Score 51; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
    |||
DB 1 TCATATAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51

RESULT 2
PCT-US01-50250A-36
; Sequence 36, Application PC/TUS0150250A
; GENERAL INFORMATION:
; APPLICANT: ELITRA PHARMACEUTICALS, INC.
; APPLICANT: HASELBECK, Robert
; APPLICANT: WALL, Daniel
; APPLICANT: GROSS, Molly
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
; FILE REFERENCE: ELITRA.010VPC
; CURRENT APPLICATION NUMBER: PCT/US01/50250A
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/259,434
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 09/948,993
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Xyl-T5 promoter sequence
PCT-US01-50250A-36

Query Match      100.0%; Score 51; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51
    |||
DB 1 TCATATAAATTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51

RESULT 3
US-10-032-393-36
; Sequence 36, Application US/10032393
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Wall, Daniel
; APPLICANT: Gross, Molly
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE
; FILE REFERENCE: ELITRA.010A
; CURRENT APPLICATION NUMBER: US/10/032,393
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/259,434
; PRIOR FILING DATE: 2000-12-27
; PRIOR APPLICATION NUMBER: 09/948,993
; PRIOR FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 51
; TYPE: DNA

```

ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Xyl-T5 promoter sequence  
US-10-032-393-36

Query Match 100.0%; Score 51; DB 43; Length 51;  
Best Local Similarity 100.0%; Pred. No. 0.00019;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTATTTGCTTCAGAGAAATTTTCTGTATATAGATTCA 51  
DB 1 TCATATAAAATTTATTTGCTTCAGAGAAATTTTCTGTATATAGATTCA 51

RESULT 4  
PCT-US01-05269-6  
Sequence 6, Application PC/TUS0105269

GENERAL INFORMATION:  
APPLICANT: The Board of Trustees of the Ireland Stanford Junior  
APPLICANT: University  
TITLE OF INVENTION: Altered Recombinases for Genome Modification  
FILE REFERENCE: 8400-0011.40  
CURRENT APPLICATION NUMBER: PCT/US01/05269  
CURRENT FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 72  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: T5, oligo 1  
PCT-US01-05269-6

Query Match 100.0%; Score 51; DB 1; Length 72;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTATTTGCTTCAGAGAAATTTTCTGTATATAGATTCA 51  
DB 2 TCATATAAAATTTATTTGCTTCAGAGAAATTTTCTGTATATAGATTCA 52

RESULT 5  
US-09-788-297-6  
Sequence 6, Application US/09788297

GENERAL INFORMATION:  
APPLICANT: Calos, Michele P.  
APPLICANT: Scilment, Christopher R.  
TITLE OF INVENTION: ALTERED RECOMBINASES FOR GENOME MODIFICATION  
FILE REFERENCE: 8400-0011  
CURRENT APPLICATION NUMBER: US/09/788,297  
CURRENT FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 72  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: T5, oligo 1  
US-09-788-297-6

Query Match 100.0%; Score 51; DB 32; Length 72;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTATTTGCTTCAGAGAAATTTTCTGTATATAGATTCA 51  
DB 2 TCATATAAAATTTATTTGCTTCAGAGAAATTTTCTGTATATAGATTCA 52

RESULT 6

PCT-US01-05269-7/C  
Sequence 7, Application PC/TUS0105269

GENERAL INFORMATION:  
APPLICANT: The Board of Trustees of the Ireland Stanford Junior  
APPLICANT: University  
TITLE OF INVENTION: Altered Recombinases for Genome Modification  
FILE REFERENCE: 8400-0011.40  
CURRENT APPLICATION NUMBER: PCT/US01/05269  
CURRENT FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 7  
LENGTH: 80  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: T5, oligo 2  
PCT-US01-05269-7

Query Match 100.0%; Score 51; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTATTTGCTTCAGAGAAATTTTCTGTATATAGATTCA 51  
DB 75 TCATATAAAATTTATTTGCTTCAGAGAAATTTTCTGTATATAGATTCA 25

RESULT 7  
PCT-US01-50250-26  
Sequence 26, Application PC/TUS0150250

GENERAL INFORMATION:  
APPLICANT: ELITRA PHARMACEUTICALS, INC.  
APPLICANT: HASELBECK, Robert  
APPLICANT: WALL, Daniel  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010VPC  
CURRENT APPLICATION NUMBER: PCT/US01/50250  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: US 60/259,434  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: US 09/948,993  
PRIOR FILING DATE: 2001-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 26  
LENGTH: 80  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Xyl-T5 fusion sequence  
PCT-US01-50250-26

Query Match 100.0%; Score 51; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTATTTGCTTCAGAGAAATTTTCTGTATATAGATTCA 51  
DB 1 TCATATAAAATTTATTTGCTTCAGAGAAATTTTCTGTATATAGATTCA 51

RESULT 8  
PCT-US01-50250A-26  
Sequence 26, Application PC/TUS0150250A

GENERAL INFORMATION:  
APPLICANT: ELITRA PHARMACEUTICALS, INC.  
APPLICANT: HASELBECK, Robert  
APPLICANT: WALL, Daniel  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010VPC

CURRENT APPLICATION NUMBER: PCT/US01/50250A  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/259,434  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: 09/948,993  
PRIOR FILING DATE: 2001-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 26  
LENGTH: 80  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Xyl-T5 fusion sequence  
PCT-US01-50250A-26

Query Match 100.0%; Score 51; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
Db 1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51

RESULT 9  
US-09-788-297-7/c  
Sequence 7, Application US/09788297  
GENERAL INFORMATION:  
APPLICANT: Calos, Michele P.  
APPLICANT: Scimemi, Christopher R.  
TITLE OF INVENTION: ALTERED RECOMBINASES FOR GENOME MODIFICATION  
FILE REFERENCE: 8400-0011  
CURRENT APPLICATION NUMBER: US/09/788,297  
CURRENT FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 7  
LENGTH: 80  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: T5, oligo 2  
US-09-788-297-7

Query Match 100.0%; Score 51; DB 32; Length 80;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
Db 75 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 25

RESULT 10  
US-10-032-393-26  
Sequence 26, Application US/10032393  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Wall, Daniel  
APPLICANT: Gross, Molly  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010A  
CURRENT APPLICATION NUMBER: US/10/032,393  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/259,434  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: 09/948,993  
PRIOR FILING DATE: 2001-09-06  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 26  
LENGTH: 80  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Xyl-T5 fusion sequence  
US-10-032-393-26

Query Match 100.0%; Score 51; DB 43; Length 80;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
Db 1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51

RESULT 11  
PCT-US01-50250-5  
Sequence 5, Application PC/TUS0150250  
GENERAL INFORMATION:  
APPLICANT: ELITRA PHARMACEUTICALS, INC.  
APPLICANT: HASSELBECK, ROBERT  
APPLICANT: WALL, DANIEL  
APPLICANT: GROSS, MOLLY  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010VPC  
CURRENT APPLICATION NUMBER: PCT/US01/50250  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: US 60/259,434  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: US 09/948,993  
PRIOR FILING DATE: 2001-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 94  
TYPE: DNA  
ORGANISM: Escherichia coli  
FEATURE:  
OTHER INFORMATION: Oligonucleotide Xyl-T5  
PCT-US01-50250-5

Query Match 100.0%; Score 51; DB 1; Length 94;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
Db 6 TCATATAAAATTATTGCTTCAGGAAATTTTCTGTATATAGATTCA 56

RESULT 12  
PCT-US01-50250A-5  
Sequence 5, Application PC/TUS0150250A  
GENERAL INFORMATION:  
APPLICANT: ELITRA PHARMACEUTICALS, INC.  
APPLICANT: HASSELBECK, ROBERT  
APPLICANT: WALL, DANIEL  
APPLICANT: GROSS, MOLLY  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010VPC  
CURRENT APPLICATION NUMBER: PCT/US01/50250A  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/259,434  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: 09/948,993  
PRIOR FILING DATE: 2001-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 94  
TYPE: DNA

ORGANISM: Escherichia coli  
FEATURE:  
OTHER INFORMATION: Oligonucleotide Xyl-T5  
PCT-US01-50250A-5

Query Match 100.0%; Score 51; DB 1; Length 94;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 6 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 56

## RESULT 13

US-10-032-393-5  
Sequence 5, Application US/10032393  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Wall, Daniel  
APPLICANT: Gross, Molly  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010A  
CURRENT APPLICATION NUMBER: US/10/032,393  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 60/259,434  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: 09/948,993  
PRIOR FILING DATE: 2001-09-06  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 94  
TYPE: DNA  
ORGANISM: Escherichia coli  
FEATURE:  
OTHER INFORMATION: Oligonucleotide Xyl-T5  
US-10-032-393-5

Query Match 100.0%; Score 51; DB 43; Length 94;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 6 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 56

## RESULT 14

US-60-259-434-3  
Sequence 3, Application US/60259434  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010PR  
CURRENT APPLICATION NUMBER: US/60/259,434  
CURRENT FILING DATE: 2000-12-27  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 94  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Artificial Oligonucleotide  
US-60-259-434-3

Query Match 100.0%; Score 51; DB 82; Length 94;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 6 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 56

## RESULT 15

PCT-US01-50250-6/c  
Sequence 6, Application PC/TUS0150250  
GENERAL INFORMATION:  
APPLICANT: ELITRA PHARMACEUTICALS, INC.  
APPLICANT: HASSELBECK, Robert  
APPLICANT: WALL, Daniel  
APPLICANT: GROSS, Molly  
TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
FILE REFERENCE: ELITRA.010VPC  
CURRENT APPLICATION NUMBER: PCT/US01/50250  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: US 60/259,434  
PRIOR FILING DATE: 2000-12-27  
PRIOR APPLICATION NUMBER: US 09/948,993  
PRIOR FILING DATE: 2001-09-06  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 95  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Oligonucleotide Xyl-T5 complement  
PCT-US01-50250-6

Query Match 100.0%; Score 51; DB 1; Length 95;  
Best Local Similarity 100.0%; Pred. No. 0.0002;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 51  
DB 93 TCATATAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGATTCA 43

Search completed: May 4, 2004, 13:15:15  
Job time : 3251 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 10:37:58 ; Search time 122 Seconds  
(without alignments)  
354.765 Million cell updates/sec

Title: US-10-032-393-36

Perfect score: 51

Sequence: 1 tcaataaaattttttgtc.....ttttctgataatgattca 51

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1490725 seqs, 424326345 residues

Total number of hits satisfying chosen parameters: 2981450

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA New: \*  
1: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq: \*  
6: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq: \*  
7: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	27.4	53.7	111469	6	US-10-767-471-10820
C 2	27	52.9	4519	6	US-10-808-570-6
C 3	26.6	52.2	7028	7	US-60-553-390-835
C 4	26	51.0	201	7	US-60-563-440-15861
C 5	26	51.0	187790	7	US-60-563-440-12163
C 6	25.6	50.2	5777	6	US-10-821-234-51
C 7	25.4	49.8	843	6	US-10-724-972A-3436
C 8	25.4	49.8	405740	6	US-10-796-307-8716
C 9	25	49.0	1634	1	PCT-US03-36895A-66
C 10	25	49.0	40000	7	US-10-767-471-10915
C 11	25	49.0	57201	7	US-60-563-440-12146
C 12	25	49.0	88457	6	US-10-796-280-12211
C 13	25	49.0	130030	6	US-10-767-471-10854
C 14	25	49.0	193757	6	US-10-767-471-10614
C 15	25	49.0	583315	7	US-60-550-051-1041
C 16	25	49.0	1790242	6	US-10-767-471-10805
C 17	24.8	48.6	61635	6	US-10-765-790-105
C 18	24.8	48.6	61635	6	US-10-765-790-105
C 19	24.6	48.2	19810	7	US-60-563-440-12277
C 20	24.6	48.2	30510	6	US-10-796-280-12229
C 21	24.6	48.2	35330	6	US-10-767-471-10625
C 22	24.6	48.2	84405	6	US-10-601-886A-31
C 23	24.6	48.2	344805	6	US-10-779-271-1
C 24	24.6	48.2	354592	6	US-10-765-790-70
C 25	24.6	48.2	1126118	6	US-10-767-471-10615
C 26	24.6	48.2	1126118	6	US-10-796-280-12479

C 27	24.6	48.2	1126118	6	US-10-796-307-8715	Sequence 8715, Ap
C 28	24.4	47.8	19053	7	US-60-563-440-12000	Sequence 12000, A
C 29	24.2	47.5	1650	6	US-10-797-366-254	Sequence 254, App
C 30	24.2	47.5	1650	6	US-10-771-187-254	Sequence 254, App
C 31	24.2	47.5	16201	6	US-10-767-471-10705	Sequence 10705, A
C 32	24.2	47.5	33231	6	US-10-765-790-45	Sequence 45, Appl
C 33	24.2	47.5	61528	7	US-60-563-440-12210	Sequence 12210, A
C 34	24.2	47.5	68196	6	US-10-417-375A-160	Sequence 160, App
C 35	24.2	47.5	68196	6	US-10-417-375B-160	Sequence 160, App
C 36	24	47.1	606	6	US-10-363-345A-37629	Sequence 37629, A
C 37	24	47.1	606	6	US-10-363-345A-37630	Sequence 37630, A
C 38	24	47.1	1200	6	US-10-767-701-13802	Sequence 13802, A
C 39	24	47.1	35126	6	US-10-796-307-8785	Sequence 8785, Ap
C 40	24	47.1	74870	6	US-10-796-280-12297	Sequence 12297, A
C 41	24	47.1	205784	7	US-60-550-051-3026	Sequence 3026, Ap
C 42	24	47.1	305312	7	US-60-550-051-2989	Sequence 2989, Ap
C 43	24	47.1	1790242	6	US-10-767-471-10805	Sequence 10805, A
C 44	23.8	46.7	201	6	US-10-767-471-16374	Sequence 16374, A
C 45	23.8	46.7	201	6	US-10-796-280-46448	Sequence 46448, A

#### ALIGNMENTS

RESULT 1  
US-10-767-471-10820/c  
Sequence 10820, Application US/10767471  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001505  
CURRENT APPLICATION NUMBER: US/10/767,471  
CURRENT FILING DATE: 2004-01-30  
NUMBER OF SEQ ID NOS: 50231  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 10820  
LENGTH: 111469  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(111469)  
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1  
US-10-767-471-10820

Query Match 53.7%; Score 27.4; DB 6; Length 111469;  
Best Local Similarity 75.6%; Pred. No. 30;  
Matches 34; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 7 AAAATTATTTGCTTCCAGAAAATTTCTGTATATGATTCA 51  
DB 39669 AATTTTCATTAGATTTCACAAAATTTTCTGTATATGATTCA 39625

RESULT 2  
US-10-808-570-6  
Sequence 6, Application US/10808570  
GENERAL INFORMATION:  
APPLICANT: PATRICK J. DILLON et al.  
TITLE OF INVENTION: Nucleotide Sequences of Escherichia coli Pathogenicity Islands  
FILE REFERENCE: PB324D1  
CURRENT APPLICATION NUMBER: US/10/808,570  
CURRENT FILING DATE: 2004-03-25  
PRIOR APPLICATION NUMBER: US/09/956,004  
PRIOR FILING DATE: 2001-09-20  
PRIOR APPLICATION NUMBER: 08/976,259  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/061,953  
PRIOR FILING DATE: 1997-10-14  
PRIOR APPLICATION NUMBER: 60/031,626  
PRIOR FILING DATE: 1996-11-22  
NUMBER OF SEQ ID NOS: 142

```
SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 4519
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURES:
;   NAME/KEY: misc feature
;   LOCATION: (3463)..(3483)
;   OTHER INFORMATION: n equals a, t, g, or c
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (3487)..(3487)
;   OTHER INFORMATION: n equals a, t, g, or c
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (4292)..(4292)
;   OTHER INFORMATION: n equals a, t, g, or c
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (4318)..(4318)
;   OTHER INFORMATION: n equals a, t, g, or c
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (4329)..(4329)
;   OTHER INFORMATION: n equals a, t, g, or c
US-10-806-570-6
```

```
Query Match          52.9%; Score 27; DB 6; Length 4519;
Best Local Similarity 70.6%; Pred. No. 37;
Matches 36; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
```

```
QY 1 TCAATTAATTTATTTCTTCAGGAAATTTTCTGTATATAGATTCA 51
DB 981 TCTTCATCAATTTTTCATTCAGGAAATTTTCAAGCAATATATATCA 1031
```

```
RESULT 3
US-60-552-390-835
; Sequence 835, Application US/60552390
; GENERAL INFORMATION:
;   APPLICANT: DOMON, Bruno
;   APPLICANT: HE, Tao
;   APPLICANT: LI, Aiqun
;   APPLICANT: ZHANG, Xiaolong
;   APPLICANT: KETCHUM, Karen
;   APPLICANT: MCCAFERY, Ian
;   APPLICANT: NARAYAN, Vaibhav
;   TITLE OF INVENTION: PANCREATIC DISEASE TARGETS AND USES
;   FILE REFERENCE: C1001514PROV
;   CURRENT APPLICATION NUMBER: US/60/552,390
;   CURRENT FILING DATE: 2004-03-12
;   NUMBER OF SEQ ID NOS: 1027
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 835
;   TYPE: DNA
;   LENGTH: 7028
;   ORGANISM: Homo sapiens
US-60-552-390-835
```

```
Query Match          52.2%; Score 26.6; DB 7; Length 7028;
Best Local Similarity 71.4%; Pred. No. 48;
Matches 35; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

```
QY 1 TCAATTAATTTATTTCTTCAGGAAATTTTCTGTATATAGATT 49
DB 5402 TGAATTAATTTATTTATTTTATGAAATTTTATATATAAAT 5450
```

```
RESULT 4
US-60-563-440-35861
; Sequence 35861, Application US/60563440
; GENERAL INFORMATION:
```

```
APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
;   FILE REFERENCE: C1001470
;   CURRENT APPLICATION NUMBER: US/60/563,440
;   CURRENT FILING DATE: 2004-04-20
;   NUMBER OF SEQ ID NOS: 47859
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 35861
;   LENGTH: 201
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-60-563-440-35861
```

```
Query Match          51.0%; Score 26; DB 7; Length 201;
Best Local Similarity 76.2%; Pred. No. 65;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 6 AAAATTTATTTCTTCAGGAAATTTTCTGTATATAGA 47
DB 33 AAAATTTATTTCTTCAGGAAATTTTATATATATA 74
```

```
RESULT 5
US-60-563-440-12163
; Sequence 12163, Application US/60563440
; GENERAL INFORMATION:
;   APPLICANT: CARGILL, Michele
;   APPLICANT: CHANG, Sheng-Yung
;   TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;   TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
;   TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
;   FILE REFERENCE: C1001470
;   CURRENT APPLICATION NUMBER: US/60/563,440
;   CURRENT FILING DATE: 2004-04-20
;   NUMBER OF SEQ ID NOS: 47859
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 12163
;   LENGTH: 187790
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (1)..(187790)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-60-563-440-12163
```

```
Query Match          51.0%; Score 26; DB 7; Length 187790;
Best Local Similarity 76.2%; Pred. No. 74;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 6 AAAATTTATTTCTTCAGGAAATTTTCTGTATATAGA 47
DB 132651 AAAATTTATTTCTTCAGGAAATTTTATATATATA 132692
```

```
RESULT 6
US-10-821-234-51/C
; Sequence 51, Application US/10821234
; GENERAL INFORMATION:
;   APPLICANT: Labat, Ivan
;   APPLICANT: Seacane-Crain, Birgit
;   APPLICANT: Andarmant, Susan
;   APPLICANT: Tang, Y. Tom
;   TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
;   FILE REFERENCE: 821A
;   CURRENT APPLICATION NUMBER: US/10/821,234
;   CURRENT FILING DATE: 2004-04-07
;   PRIOR APPLICATION NUMBER: US 60/462,047
```



PRIOR FILING DATE: 2003-04-07  
NUMBER OF SEQ ID NOS: 1704  
SOFTWARE: PL\_SEQ\_genes Version 1.0  
SEQ ID NO 51  
LENGTH: 5777  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-821-234-51

Query Match 50.2%; Score 25.6; DB 6; Length 5777;  
Best Local Similarity 70.8%; Pred. No. 90;  
Matches 34; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTTGGCTTCAGGAAATTTTCTGTATATAGAT 48  
DB 5463 TTATATAAACTTCAATTGAAATGACAAAGTTCTTTTAAATGAT 5416

## RESULT 7

US-10-724-972A-3436/c  
Sequence 3436, Application US/10724972A  
GENERAL INFORMATION:  
APPLICANT: Doucette-Stamm, Lynn  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: PAT03-16  
CURRENT APPLICATION NUMBER: US/10/724,972A  
PRIOR FILING DATE: 2003-12-01, 969  
PRIOR APPLICATION NUMBER: 09/450,969  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: 09/134,001  
PRIOR FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 7544  
SEQ ID NO 3436  
LENGTH: 843  
TYPE: DNA  
ORGANISM: S.epidermidis  
US-10-724-972A-3436

Query Match 49.8%; Score 25.4; DB 6; Length 843;  
Best Local Similarity 68.6%; Pred. No. 99;  
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTTGGCTTCAGGAAATTTTCTGTATATAGATCA 51  
DB 476 TAATCATACATATATGCTCTTCAGCAATTTTGTCAATATAGCCTCA 426

## RESULT 8

US-10-796-307-8716  
Sequence 8716, Application US/10796307  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: C1001509  
CURRENT APPLICATION NUMBER: US/10/796,307  
CURRENT FILING DATE: 2004-03-10  
NUMBER OF SEQ ID NOS: 44201  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8716  
LENGTH: 405740  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-796-307-8716

Query Match 49.8%; Score 25.4; DB 6; Length 405740;  
Best Local Similarity 68.6%; Pred. No. 1.1e+02;

Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTTGGCTTCAGGAAATTTTCTGTATATAGATCA 51  
DB 141784 TCACAAATCATTTAGTTATTTTACTTAACATCTCTGTATATGCTCA 141834

## RESULT 9

PCT-US03-36895A-66/c  
Sequence 66, Application PC/TUS0336895A  
GENERAL INFORMATION:  
APPLICANT: MONSANTO TECHNOLOGY LLC  
APPLICANT: Ruzinsky, Diane  
APPLICANT: Benetec, Kristen  
APPLICANT: Jander, Georg  
TITLE OF INVENTION: Production of Increased Oil and Protein in Plants by the  
FILE REFERENCE: REN-01-024  
CURRENT APPLICATION NUMBER: PCT/US03/36895A  
CURRENT FILING DATE: 2003-11-17  
PRIOR APPLICATION NUMBER: 60/427,313  
PRIOR FILING DATE: 2002-11-18  
NUMBER OF SEQ ID NOS: 176  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 66  
LENGTH: 1634  
TYPE: DNA  
ORGANISM: Glycine max  
PCT-US03-36895A-66

Query Match 49.0%; Score 25; DB 1; Length 1634;  
Best Local Similarity 69.4%; Pred. No. 1.3e+02;  
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTTGGCTTCAGGAAATTTTCTGTATATAGAT 49  
DB 1610 TCATATAAAATTTTGGCTTCAGGAAATTTTGTATCATGAAAT 1562

## RESULT 10

US-10-767-471-10915/c  
Sequence 10915, Application US/10767471  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: C1001505  
CURRENT APPLICATION NUMBER: US/10/767,471  
CURRENT FILING DATE: 2004-01-30  
NUMBER OF SEQ ID NOS: 50231  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10915  
LENGTH: 40000  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(40000)  
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1  
US-10-767-471-10915

Query Match 49.0%; Score 25; DB 6; Length 40000;  
Best Local Similarity 69.4%; Pred. No. 1.4e+02;  
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 TCATATAAAATTTTGGCTTCAGGAAATTTTCTGTATATAGAT 49  
DB 36399 TAAGAAAAATATATTTTCTATTAAGCACTATTTAGAGTAATAT 36351

## RESULT 11

US-60-563-440-12146/c  
Sequence 12146, Application US/60563440

GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C  
TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES  
FILE REFERENCE: CL001470  
CURRENT APPLICATION NUMBER: US/60/563,440  
NUMBER OF SEQ ID NOS: 47859  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12146  
LENGTH: 57201  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-60-563-440-12146

Query Match 49.0%; Score 25; DB 7; Length 57201;  
Best Local Similarity 75.6%; Pred. No. 1.4e+02;  
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Cy 4 TAAAAAATTATTTGCTTCAGGAAATTTTCTGTATTAAT 44  
DB 23118 TAAATGATGATTCCTTCAGGAAATTTATCTCTATT 23078

RESULT 12  
US-10-796-280-12211  
Sequence 12211, Application US/10796280

GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001510  
CURRENT APPLICATION NUMBER: US/10/796,280  
NUMBER OF SEQ ID NOS: 68533  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12211  
LENGTH: 88457  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-796-280-12211

Query Match 49.0%; Score 25; DB 6; Length 88457;  
Best Local Similarity 69.4%; Pred. No. 1.4e+02;  
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Cy 1 TCATAAAAATTATTTGCTTCAGGAAATTTTCTGTATTAATAGATT 49  
DB 31986 TCTTAGAATTTATTTCTTCATGCAATATCTTATCTATT 32034

RESULT 13  
US-10-767-471-10854  
Sequence 10854, Application US/10767471

GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001505  
CURRENT APPLICATION NUMBER: US/10/767,471  
NUMBER OF SEQ ID NOS: 50231  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10854  
LENGTH: 130030  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-767-471-10854

Query Match 49.0%; Score 25; DB 6; Length 130030;

Best Local Similarity 69.4%; Pred. No. 1.4e+02;  
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Cy 3 ATAAAAATTATTTGCTTCAGGAAATTTTCTGTATTAATAGATTCA 51  
DB 79661 ATAAATAAATATTTCTTCCTTATGATTTTCTTAATATATTTTCA 79709

RESULT 14  
US-10-767-471-10614/c  
Sequence 10614, Application US/10767471

GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001505  
CURRENT APPLICATION NUMBER: US/10/767,471  
NUMBER OF SEQ ID NOS: 50231  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10614  
LENGTH: 193757  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-767-471-10614

Query Match 49.0%; Score 25; DB 6; Length 193757;  
Best Local Similarity 69.4%; Pred. No. 1.4e+02;  
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Cy 3 ATAAAAATTATTTGCTTCAGGAAATTTTCTGTATTAATAGATTCA 51  
DB 27836 ATAAATAAATATTTCTTCCTTATGATTTTCTTAATATATTTTCA 27788

RESULT 15  
US-60-550-051-3041  
Sequence 3041, Application US/60550051

GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001508  
CURRENT APPLICATION NUMBER: US/60/550,051  
NUMBER OF SEQ ID NOS: 23014  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3041  
LENGTH: 583315  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(583315)  
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (See Tables 1  
US-60-550-051-3041

Query Match 49.0%; Score 25; DB 7; Length 583315;  
Best Local Similarity 75.6%; Pred. No. 1.4e+02;  
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Cy 4 TAAAAAATTATTTGCTTCAGGAAATTTTCTGTATTAAT 44  
DB 352670 TACCAATTTATTTCTTCAGGAAATTTTCTTATT 352710

Search completed: May 4, 2004, 13:17:32  
Job time : 132 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 09:47:38 ; Search time 2695 Seconds

(without alignments)  
565.110 Million cell updates/sec

Title: US-10-032-393-36

Perfect score: 51

Sequence: 1 tcaataaaatttatttgc.....ttttcgtataatgatcca 51

Scoring table:

IDENTITY NUC  
Gapop 10-0, Gapex 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:  
1: em\_estda:  
2: em\_esthm:  
3: em\_estlm:  
4: em\_estmu:  
5: em\_estov:  
6: em\_estp:  
7: em\_estro:  
8: em\_estc:  
9: gb\_estc1:  
10: gb\_estc2:  
11: gb\_estc3:  
12: gb\_estc4:  
13: gb\_estc5:  
14: gb\_estc6:  
15: em\_estlm:  
16: em\_estom:  
17: em\_gss\_hum:  
18: em\_gss\_inv:  
19: em\_gss\_pln:  
20: em\_gss\_vit:  
21: em\_gss\_fun:  
22: em\_gss\_mam:  
23: em\_gss\_mus:  
24: em\_gss\_pro:  
25: em\_gss\_rnd:  
26: em\_gss\_phg:  
27: em\_gss\_vr1:  
28: gb\_gss1:  
29: gb\_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30.2	59.2	318	14	CF350907
2	28.4	57.6	358	9	AM049113
3	28.4	57.6	534	10	BB365884
4	28.4	57.6	667	10	BB080413

Result No.	Score	Query Match	Length	ID	Description
5	29.4	57.6	768	13	BO174160
6	29.4	57.6	2505	11	AK048039
7	29.4	57.6	3403	11	AK034193
8	29.2	57.3	356	9	A1394884
9	29.2	57.3	412	9	AA219512
10	29.2	57.3	130	28	B2128980
11	29	56.9	595	28	B2145989
12	29	56.9	1048	28	B2184761
13	28.8	56.5	560	28	B2194126
14	28.8	56.5	654	28	B2216912
15	28.8	56.5	773	14	CF406767
16	28.8	56.5	875	28	B2192372
17	28.4	55.7	412	29	AG242333
18	28.4	55.3	140	28	AF056259
19	28.2	55.3	611	28	BH741163
20	28.2	55.3	733	28	BH957079
21	28.2	55.3	757	28	BH992006
22	28.2	55.3	834	13	BUT46323
23	28	54.9	659	14	CB431012
24	28	54.9	729	14	CB430297
25	27.8	54.5	381	10	AM522148
26	27.8	54.5	748	28	BH670666
27	27.6	54.1	438	28	AO782840
28	27.6	54.1	824	28	BH669814
29	27.4	53.7	571	28	BH506288
30	27.4	53.7	603	9	AA799484
31	27.4	53.7	641	28	B2082391
32	27.4	53.7	702	29	CE003137
33	27.4	53.7	718	28	BH995398
34	27.4	53.7	760	28	BH605836
35	27.4	53.7	993	28	BH692567
36	27.4	53.7	1340	29	CG751618
37	27.2	53.3	101	28	B2217382
38	27.2	53.3	292	9	AA365198
39	27.2	53.3	357	28	B2268137
40	27.2	53.3	378	28	B2199225
41	27.2	53.3	450	29	CE706869
42	27.2	53.3	544	28	B2217207
43	27.2	53.3	546	28	B2177519
44	27.2	53.3	560	28	B2214923
45	27.2	53.3	631	28	B2219091

## ALIGNMENTS

RESULT 1  
CF350907  
LOCUS  
DEFINITION  
CDNA 5', mRNA sequence.  
ACCESSION  
CF350907.1 GI:33953420  
VERSION  
CF350907.1  
KEYWORDS  
SOURCE  
ORGANISM  
Meloidogyne javanica (root-knot nematode)  
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina; Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.  
REFERENCE  
1 (bases 1 to 318)  
McCarter, J., Clifton, S., Chiappelli, B., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritzer, E., Bennett, J., Franklin, C., Tsagaris, V., R., Konko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Stepien, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterson, R. and Wilson, R.  
The Washington Univ. Nematode EST Project, 1999  
Unpublished (1999)  
Contact: McCarter JP  
The Washington Univ. Nematode EST Project, 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800

Fax: 314 286 1810  
Email: est@atascn.wustl.edu  
Cloned unidirectionally. Poly(A) + RNA was concentrated and purified using Dynabeads (Dyna) and mRNA eluted for first strand synthesis. First strand cDNA was created using MMLV RT (Powerscript, Clontech) and primed with oligo(dT) with XhoI site and 5'SMART 'anchor' added using chimeric DNA-RNA oligo. 12 PCR cycles were done using first strand and primers specific to SMART oligo and 3' end. Double stranded cDNA was digested using XhoI/NotI, fractionated on Chroma-spin 400 columns (Clontech) and ligated to digested pGEM-11zf(+) plasmid. Chemically competent DH10B cells were used as host cells. Library materials provided by Dr. David Bird of North Carolina State University. Library construction by Jeff Rousch. See www.nematode.net for additional project information.  
The vector to vector length is 319  
Seq primer: Sp6.

## FEATURES

## source

Location/Qualifiers  
1..318  
/organism="Meloidogyne javanica"  
/mol\_type="mRNA"  
/db\_xref="taxon:6303"  
/tissue\_type="whole organism"  
/dev\_stage="J2"  
/lab\_host="DH10B"  
/clone\_lib="Meloidogyne javanica J2 SMART pGEM"  
/note="Vector: plasmid (ampicillin resistant); Site 1: XhoI; Site 2: NotI; Cloned unidirectionally. Poly(A) + RNA was concentrated and purified using Dynabeads (Dyna) and mRNA eluted for first strand synthesis. First strand cDNA was created using MMLV RT (Powerscript, Clontech) and primed with oligo(dT) with XhoI site and 5'SMART 'anchor' added using chimeric DNA-RNA oligo. 12 PCR cycles were done using first strand and primers specific to SMART oligo and 3' end. Double stranded cDNA was digested using XhoI/NotI, fractionated on Chroma-spin 400 columns (Clontech) and ligated to digested pGEM-11zf(+) plasmid. Chemically competent DH10B cells were used as host cells. Library materials provided by Dr. David Bird of North Carolina State University. Library construction by Jeff Rousch. See www.nematode.net for additional project information."

## ORIGIN

Query Match 59.2%; Score 30.2; DB 14; Length 318;  
Best Local Similarity 81.4%; Pred. No. 2.9e+02;  
Matches 35; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 CATAAAATTTATTTGCTTCAGGAAATTTTCTGTATAT 44  
80 CAAAAGAAATTTATTTATATCAGGAAATTTTCTGTATAT 122

RESULT 2  
LOCUS AW049113 358 bp mRNA linear EST 18-SEP-1999  
DEFINITION UI-M-BH1-ant-g-11-0-UI .s1 NIH BMAP M S2 Mus musculus cDNA clone  
ACCESSION AW049113  
VERSION AW049113.1 GI:5909642  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 358)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene discovery  
Genome Res. 6 (9), 791-806 (1996)  
JOURNAL 97044477  
MEDLINE 8889548  
COMMENT Contact: Chin, H

## FEATURES

## source

Location/Qualifiers  
1..358  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone\_lib="UI-M-BH1-ant-g-11-0-UI"  
/dev\_stage="27-32 days"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH-BMAP M S2 library is a subtracted library derived from NIH-BMAP M S1, which in turn is a subtracted library derived from a mixture of normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus). The driver used for subtraction consisted of a pool of 5,000 clones from the NIH-BMAP M S1 library and a pool of 2,000 clones obtained from non-normalized and normalized mouse brain spinal cord libraries.  
TAG TRISUE=basal-ganglia  
TAG\_LIB=NIH\_BMAP\_M\_S2  
TAG\_SEQ=GTGAC"

## ORIGIN

Query Match 57.6%; Score 29.4; DB 9; Length 358;  
Best Local Similarity 76.6%; Pred. No. 4.5e+02;  
Matches 36; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 3 ATAAATTTATTTGCTTCAGGAAATTTTCTGTATATAGAT 49  
20 ATAAATTTATTTGCTTCAGGAAATTTTCTGTATATACAT 66

RESULT 3  
LOCUS BB365884/c 534 bp mRNA linear EST 24-OCT-2001  
DEFINITION BB365884 RIKEN full-length enriched, 16 days embryo head Mus musculus cDNA clone C13031F03 3', mRNA sequence.  
ACCESSION BB365884  
VERSION BB365884.2 GI:16406384  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 534)  
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Komno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Onno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,

TITLE  
JOURNAL  
COMMENT

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M., and Hayashizaki, Y. RIKEN Mouse ESTs (Arakawa, T., et al. 2001) Unpublished (2001)

On Jul 12, 2000 this sequence version replaced gi:9077712. Contact: Yoshinide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suenho-cho, Tsunumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216

FEATURES  
SOURCE

Email: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamana, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., and Hayashizaki, Y. Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Location/Qualifiers 1. 534 /organism="Mus musculus" /mol\_type="mRNA" /strain="C57BL/6J" /db\_xref="taxon:10090" /clone="C130031P03" /sex="mixed" /tissue\_type="head" /dev\_stage="16 days embryo" /lab\_host="DH10B" /clone\_lib="RIKEN full-length enriched, 16 days embryo head" /note="Site 1: Salt, Site 2: BamHI, cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGAGTTATTAATTAATCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from lambda FLC I"

## ORIGIN

Query Match 57.6%; Score 29.4; DB 10; Length 534; Best Local Similarity 76.6%; Pred. No. 3.9e+02; Matches 36; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

3 ATAAAAATTTATTTGCTTTCAGAAATTTTCTGTAATTAAGATT 49  
ATTATTTATTTATTTGCTTTCAGAAATTTTCTGTAATTAAGATT 259

BB080413 667 bp mRNA linear EST 18-OCT-2001  
BB080413 RIKEN full-length enriched, adult male diencephalon Mus musculus cDNA clone 9330162L04 3', mRNA sequence.

BB080413.2 GI:16260594  
EST.

Mus musculus (house mouse)

## ORGANISM

## REFERENCE

## AUTHORS

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Kono, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, K., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shitoh, R., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M., and Hayashizaki, Y. RIKEN Mouse ESTs (Arakawa, T., et al. 2001) Unpublished (2001)

## TITLE

## JOURNAL

## COMMENT

Contract: Yoshinide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suenho-cho, Tsunumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamana, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., and Hayashizaki, Y. Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Location/Qualifiers

## FEATURES



Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, U., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE 20530913 11076861

PUBMED

REFERENCES

AUTHORS

4

THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE

Functional annotation of a full-length mouse cDNA collection

JOURNAL NATURE 409, 685-690 (2001)

REFERENCE

AUTHORS

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL NATURE 420, 563-573 (2002)

REFERENCE

AUTHORS

6 (bases 1 to 2505)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Muramatsu, M. and Hayashizaki, Y.

Direct Submission

TITLE

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/.

Location/Qualifiers

FEATURES

SOURCE

1. 2505

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="RANTOM:DB:CL30031F03"

/db\_xref="MGI:2414068"

/db\_xref="taxon:10090"

/clone="CL30031F03"

/issue\_type="head"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="16 days embryo"

1. 2505

/note="unknown EST (GB|B173511, evidence: BLASTN, 99%, match=817)"

misc\_feature

ORIGIN

Query Match 57.6%; Score 29.4; DB 11; Length 2505;

Best Local Similarity 76.6%; Pred. No. 2.3e+02;

Matches 36; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

3 ATAAATTTATTTGCTTTCAGGAAATTTTCTGTAATATAGATT 49

2274 ATAAATTTATTTGCTTTCAGGAAATTTTCTGTAATATAGATT 2228

RESULT 7

AK034193/c

LOCUS

DEFINITION

AK034193 3403 bp mRNA linear HTC 18-SEP-2003

Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched library, clone:9330162L04 product:unknown EST, full insert sequence.

ACCESSION

AK034193

VERSION

AK034193.1 GI:26083816

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

1

Carninci, P. and Hayashizaki, Y.

TITLE

High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE

AUTHORS

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE

AUTHORS

3

Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Teshiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, U., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE 20530913 11076861

PUBMED

REFERENCES

AUTHORS

4

THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE

Functional annotation of a full-length mouse cDNA collection

JOURNAL NATURE 409, 685-690 (2001)

REFERENCE

AUTHORS

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL NATURE 420, 563-573 (2002)

REFERENCE

AUTHORS

6 (bases 1 to 3403)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Muramatsu, M. and Hayashizaki, Y.

Direct Submission

TITLE

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT Fax:81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome

Enyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in Riken.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: <http://genome.gsc.riken.go.jp/>  
URL: <http://fantom.gsc.riken.go.jp/>  
Location/Qualifiers

## FEATURES

source

1..3403  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM,DB:9330162L04"  
/db\_xref="MGI:2398388"  
/db\_xref="taxon:10090"  
/clone="9330162L04"  
/sex="male"  
/tissue\_type="diencephalon"  
/clone\_id="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
1..3403  
/note="unknown EST (GB|B1735511, evidence: BLASTN, 99%,  
match=817)"

## ORIGIN

Query Match 57.6% Score 29.4; DB 11; Length 3403;  
Best Local Similarity 76.6%; Pred. No. 2.1e+02;  
Matches 36; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 3 ATAAATTTATTTGCTTCAGGAAATTTTCTGTATATACATT 49

Db 3391 ATATATTTTATTTGCTTCAGGAAATTTTCTGTATATACATT 3345

RESULT 8 A1394884 356 bp mRNA linear EST 04-FEB-1999  
LOCUS A1394884  
DEFINITION MA001291.C8F Soares normalized S8W Schistosoma mansoni CDNA 3',  
mRNA sequence.

ACCESSION A1394884  
VERSION A1394884.1 GI:4224431  
KEYWORDS EST.

SOURCE Schistosoma mansoni

ORGANISM Schistosoma mansoni

REFERENCE Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;  
Strigoididae; Schistosomatidae; Schistosomatidae; Schistosoma.

AUTHORS Bailey, J.A., Bentley, K., Blanton, R.E., Soares, M.B. and  
Chakravarti, A.

TITLE Expressed sequence tags from a normalized adult Schistosoma mansoni  
library

JOURNAL Unpublished (1999)

COMMENT Contact: Chakravarti A  
Department of Genetics  
Case Western Reserve University  
2109 Adelbert Rd, Cleveland, OH 44106, USA  
Tel: 216-368-5847  
Fax: 216-368-5857  
Email: axc39@po.cwrn.edu  
Additional data regarding this EST may be found at  
<http://genome.cwrn.edu/schistosoma/est/S8west.html>  
Seq primer: M13 Forward.

FEATURES Location/Qualifiers

source

1..356  
/organism="Schistosoma mansoni"  
/mol\_type="mRNA"  
/strain="Puerto Rican"  
/db\_xref="taxon:6183"  
/sex="mixed"  
/dev\_stage="8 week old adult worms"  
/lab\_host="Mus musculus"

## ORIGIN

Query Match 57.3% Score 29.2; DB 9; Length 356;  
Best Local Similarity 74.0%; Pred. No. 5.1e+02;  
Matches 37; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CATTAATTTATTTGCTTCAGGAAATTTTCTGTATATACATTCA 51

Db 140 CATTAATTTATTTGCTTCAGGAAATTTTCTGTATATACATTCA 189

RESULT 9 AA219512 412 bp mRNA linear EST 07-FEB-1997  
LOCUS AA219512/c  
DEFINITION z999G04.r1 Striatogene NT2 neuronal precursor 937230 Homo sapiens  
CDNA clone IMAGE:650166 5', mRNA sequence.

ACCESSION AA219512  
VERSION AA219512.1 GI:1833571  
KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,  
Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,  
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,  
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
Rohtling, T., Scheinberg, K., Soares, M.B., Tan, F., Thierry-Veg, J.,  
Trevaaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.  
and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags

JOURNAL Genome Res. 6 (9), 807-828 (1996)

MEDLINE 97044478

PUBMED 8889549

COMMENT Contact: Wilson RK

Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)

This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Seq primer: 28m13 rev1 ET from Amersham  
High quality sequence Btop: 262.

## FEATURES

source

1..412  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:5276755"  
/db\_xref="taxon:9606"  
/clone="IMAGE:650166"  
/tissue\_type="neuroepithelial cells"  
/dev\_stage="NTera-2 neuroepithelial cells"  
/lab\_host="SOLR (kanamycin resistant)"  
/clone\_id="Striatogene NT2 neuronal precursor 937230"  
/note="Organ: brain; Vector: pBluescript SK-; Site: 1;  
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:  
Oligo dt. Uninduced, exponentially growing neuroepithelial  
cells (NTera-2/cl.D1). Average insert size: 1.0 kb;  
Uni-ZAP XR Vector); ~5' adaptor sequence: 5' GAATTGGGCGAG  
3' ~3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"

## ORIGIN

Query Match

57.3%; Score 29.2; DB 9; Length 412;



Best Local Similarity 72.5%; Pred. No. 4.8e+02;  
Matches 37; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCATAAAATTTATTTGCTTCAGGAAATTTTCTGTATATGATTCA 51  
122 TCATATAAAAGATGATTTCTTCTTCTTAAATTTCTGTATCTAAATTTA 72

RESULT 10 BZ128980 130 bp DNA linear GSS 11-OCT-2002  
LOCUS BZ128980  
DEFINITION CH230-452M22.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone  
CH230-452M22, genomic survey sequence.

ACCESSION BZ128980  
VERSION BZ128980.1 GI:23769927  
KEYWORDS GSS.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 130)  
AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,  
Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,  
Riggs, F., de Jong, P. and Fraser, C.M.  
Rat BAC End Sequences from Library CHORI-230 MboI segment  
Unpublished (1999)  
COMMENT Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230  
(http://www.chori.org/bacpac/rat230.htm). For BAC library  
availability, please contact Pieter de Jong (pdejong@mail.chori.org).  
Clones may be purchased from BACPAC Resources  
(http://www.chori.org/bacpac/or ering information.htm). BAC end  
page: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html  
Plate: 452 row: M column: 22  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source location/Qualifiers

1..130  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/strain="BN/SsNHsd/MCW"  
/db\_xref="taxon:10116"  
/clone="CH230-452M22"  
/sex="Female"  
/cell\_type="Brain"  
/clone\_id="CHORI-230 Segment 2"  
/note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;  
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by  
Pieter de Jong"

ORIGIN  
Query Match 56.9%; Score 29; DB 28; Length 130;  
Best Local Similarity 86.5%; Pred. No. 8e+02;  
Matches 32; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 AAAAATTTATTTGCTTCAGGAAATTTTCTGTAT 41  
10 AAAAATTTGATTCATTCATGAAATTTTCTGTAT 46

RESULT 11 BZ145989 595 bp DNA linear GSS 11-OCT-2002  
LOCUS BZ145989  
DEFINITION CH230-452M22.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone  
CH230-452M22, genomic survey sequence.

VERSION BZ145989.1 GI:23766936  
KEYWORDS GSS.  
SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 595)  
AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,  
Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,  
Riggs, F., de Jong, P. and Fraser, C.M.  
Rat BAC End Sequences from Library CHORI-230 MboI segment  
Unpublished (1999)  
COMMENT Other GSSs: CH230-452M22.TJ  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230  
(http://www.chori.org/bacpac/rat230.htm). For BAC library  
availability, please contact Pieter de Jong (pdejong@mail.chori.org).  
Clones may be purchased from BACPAC Resources  
(http://www.chori.org/bacpac/or ering information.htm). BAC end  
page: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html  
Plate: 452 row: N column: 22  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source location/Qualifiers

1..595  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/strain="BN/SsNHsd/MCW"  
/db\_xref="taxon:10116"  
/clone="CH230-452M22"  
/sex="Female"  
/cell\_type="Brain"  
/clone\_id="CHORI-230 Segment 2"  
/note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;  
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by  
Pieter de Jong"

ORIGIN  
Query Match 56.9%; Score 29; DB 28; Length 595;  
Best Local Similarity 86.5%; Pred. No. 4.8e+02;  
Matches 32; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 AAAAATTTATTTGCTTCAGGAAATTTTCTGTAT 41  
465 AAAAATTTGATTCATTCATGAAATTTTCTGTAT 521

RESULT 12 BZ184761 1048 bp DNA linear GSS 11-OCT-2002  
LOCUS BZ184761  
DEFINITION CH230-423F16.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone  
CH230-423F16, genomic survey sequence.

ACCESSION BZ184761  
VERSION BZ184761.1 GI:23834700  
KEYWORDS GSS.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 1048)  
AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,  
Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,  
Riggs, F., de Jong, P. and Fraser, C.M.  
Rat BAC End Sequences from Library CHORI-230 MboI segment  
Unpublished (1999)

## COMMENT

Other GSSs: CH230-423F16.TV  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208

Email: szhao@tigr.org  
 Clones are derived from the rat BAC library CHORI-230  
 (http://www.chori.org/bacpac/rat230.htm). For BAC library  
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
 Clones may be purchased from BACPAC Resources  
 (http://www.chori.org/bacpac/orering\_information.htm). BAC end  
 page: http://www.tigr.org/tcd/bac\_ends/rat/bac\_end\_intro.html  
 Plate: 423 row: F column: 16  
 Seq primer: SP6  
 Class: BAC ends.

## FEATURES

## source

Location/Qualifiers  
 1..1048

/organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /strain="BN/SMNHd/MCW"  
 /db\_xref="taxon:10116"  
 /clone="CH230-423F16"  
 /sex="Female"  
 /cell\_type="Brain"  
 /clone\_jib="CHORI-230 Segment 2"  
 /note="Vector: pTARBA1.3; Site 1: Mbol; Site 2: Mbol;  
 CHORI-230 Rat (BN/SMNHd/MCW) BAC library produced by  
 Pieter de Jong"

## ORIGIN

Query Match 56.9%; Score 29; DB 28; Length 1048;  
 Best Local Similarity 86.5%; Pred. No. 4e+02;  
 Matches 32; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 AAAAAAAAAATTGCTTCAGGAAAAATTTTCTGTAT 41  
 DB 108 AAAAAAAAAATTGATTCATTCATGAAAAATTTTCTGTAT 72

RESULT 13  
 BZ194126/c 560 bp DNA linear GSS 11-OCT-2002  
 LOCUS CH230-465D15.TVB CHORI-230 Segment 2 Rattus norvegicus genomic  
 DEFINITION clone CH230-465D15, genomic survey sequence.  
 accession BZ194126  
 version BZ194126.1 GI:23852178  
 keywords GSS.  
 source Rattus norvegicus (Norway rat)  
 organism Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 560)  
 Reference Zhao, S., Shetty, U., Shatsman, S., Tsagaye, G., Geer, K.,  
 Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,  
 Riggs, F., de Jong, P. and Fraser, C.M.  
 Rat BAC End Sequences from Library CHORI-230 Mbol segment  
 unpublished (1999)  
 title Rat BAC End Sequences from Library CHORI-230 Mbol segment  
 journal Other\_GSSs: CH230-465D15.TVB  
 comment Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org  
 Clones are derived from the rat BAC library CHORI-230  
 (http://www.chori.org/bacpac/rat230.htm). For BAC library  
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
 Clones may be purchased from BACPAC Resources  
 (http://www.chori.org/bacpac/orering\_information.htm). BAC end  
 page: http://www.tigr.org/tcd/bac\_ends/rat/bac\_end\_intro.html  
 Plate: 417 row: O column: 3  
 Seq primer: SP6  
 Class: BAC ends.

## FEATURES

## source

page: http://www.tigr.org/tcd/bac\_ends/rat/bac\_end\_intro.html  
 Plate: 465 row: D column: 15  
 Seq primer: SP6  
 Class: BAC ends.

## ORIGIN

Query Match 56.5%; Score 28.8; DB 28; Length 560;  
 Best Local Similarity 82.5%; Pred. No. 5.5e+02;  
 Matches 33; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CATAAAAATTATTTGCTTCAGGAAAAATTTTCTGTAT 41  
 DB 526 CATTAAAAATTATTCATTCATGAAAAATTTTCTGTAT 487

RESULT 14  
 BZ216912/c 654 bp DNA linear GSS 11-OCT-2002  
 LOCUS CH230-41703.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone  
 DEFINITION CH230-41703, genomic survey sequence.  
 accession BZ216912  
 version BZ216912.1 GI:23875270  
 keywords GSS.  
 source Rattus norvegicus (Norway rat)  
 organism Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 654)  
 Reference Zhao, S., Shetty, U., Shatsman, S., Tsagaye, G., Geer, K.,  
 Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,  
 Riggs, F., de Jong, P. and Fraser, C.M.  
 Rat BAC End Sequences from Library CHORI-230 Mbol segment  
 unpublished (1999)  
 title Rat BAC End Sequences from Library CHORI-230 Mbol segment  
 journal Other\_GSSs: CH230-41703.TV  
 comment Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org  
 Clones are derived from the rat BAC library CHORI-230  
 (http://www.chori.org/bacpac/rat230.htm). For BAC library  
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
 Clones may be purchased from BACPAC Resources  
 (http://www.chori.org/bacpac/orering\_information.htm). BAC end  
 page: http://www.tigr.org/tcd/bac\_ends/rat/bac\_end\_intro.html  
 Plate: 417 row: O column: 3  
 Seq primer: SP6  
 Class: BAC ends.

## FEATURES

## source

Location/Qualifiers  
 1..654  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /strain="BN/SMNHd/MCW"  
 /db\_xref="taxon:10116"  
 /clone="CH230-41703"  
 /sex="Female"  
 /cell\_type="Brain"

Job time : 2708 secs

/clone\_1lb="CHORI-230 Segment 2"  
 /note="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;  
 CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by  
 Pieter de Jong"

## ORIGIN

Query Match 56.5%; Score 28.8; DB 28; Length 654;  
 Best Local Similarity 82.5%; Pred. No. 5.2e+02;  
 Matches 33; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CATAAAAATTATTTGCTTCAGGAAATTTTCTGTAT 41  
 DB 513 CTTAGAAATTTTATTCATTCATGAAATTTTCTGTAT 474

## RESULT 15

CF406767/c

CF406767

773 bp mRNA linear EST 02-SEP-2003

LOCUS

CH3#044\_C08T3 Canine heart normalized cDNA library in pBluescript

DEFINITION

Canis familiaris cDNA clone CH3#044\_C08 3', mRNA sequence.

ACCESSION

VERSION

CF406767.1 GI:34407011

KEYWORDS

EST.

SOURCE

Canis familiaris (dog)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

## ORIGIN

Query Match 56.5%; Score 28.8; DB 14; Length 773;  
 Best Local Similarity 75.0%; Pred. No. 4.2e+02;  
 Matches 36; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 CATAAAAATTATTTGCTTCAGGAAATTTTCTGTATGATGAT 49  
 DB 238 CTTAAAAATTTTATTCATTCATGAAATTTTCTGTATGATGAT 191